DOCUMENT RESUME

ED 194 576 TM 800 705

AUTHOR Ward, Barbara; And Others
TITLE Major Informational Needs of National Assessment Audiences and Ways to Enhance the Assessment's Utility in Meeting Those Needs.
INSTITUTION Education Commission of the States, Denver, Colo. National Assessment of Educational Progress.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO NAEP-12-IP-51
PUB DATE Jun 80
GRANT NIE-G-80-0003
NOTE 103p.

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS *Audiences; *Educational Assessment; Elementary Secondary Education; Federal Programs; Information Dissemination; *Information Needs; *National Surveys; *Needs Assessment; Program Costs; Program Design; Program Improvement; Test Results; *User Satisfaction (Information)
IDENTIFIERS *National Assessment of Educational Progress

ABSTRACT

Audiences were surveyed at the federal, state, and local levels in an effort to determine how the National Assessment of Educational Progress could be made more useful in meeting their needs. Recommendations fell into several major categories: (1) increased dissemination of assessment materials and access to data; (2) increased data interpretation and data analysis; (3) increased collection of data on background and program variables; (4) increased research potential; (5) increased policy relevance; and (6) increased attention to state and local needs. Needs expressed by different audiences were at times in conflict. Education researchers typically would like more background and program information, while national curriculum organizations would like more achievement questions. State and local educators would like more items released for their use, while obtaining reliable change data demands that as many items as possible be kept secure. The present Assessment design does not permit some of the desired uses of assessment data. Trade-offs between various audience needs are outlined and relative costs of different procedures are indicated. (Author/RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED194576

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

MAJOR INFORMATIONAL NEEDS OF NATIONAL ASSESSMENT AUDIENCES AND WAYS TO ENHANCE THE ASSESSMENT'S UTILITY IN MEETING THOSE NEEDS

No. 12-IP-51

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
T. Pratt

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

by Barbara Ward

With Jack Schmidt, Vern Achtermann and the Staff of the
National Assessment of Educational Progress

Education Commission of the States
Suite 700, 1860 Lincoln Street
Denver, Colorado 80295

June 1980

TM 800 705

The National Assessment of Educational Progress is funded by the National Institute of Education. It is under contract with the Education Commission of the States. It is the policy of the Education Commission of the States to take affirmative action to prevent discrimination in its policies, programs and employment practices.

The National Assessment of Educational Progress is an education research project mandated by Congress to collect and report data, over time, on the performance of young Americans in various learning areas. National Assessment makes available information on assessment procedures and materials to state and local education agencies and others.

The work upon which this publication is based was performed pursuant to Grant No. NIE-G-80-0003 of the National Institute of Education. It does not, however, necessarily reflect the views of that agency.

TABLE OF CONTENTS

PREFACE. 1

INTRODUCTION 3
Utility of the National Assessment 5
Survey of Assessment Uses. 7

SECTION I Needs for, Utilization of and Suggested Enhancements to the National Assessment of Educational Progress 9
A. Needs for Assessment Data at the National Level 9
Federal Legislative Branch 10
Federal Agencies 15
National Education Organizations 20
National Curriculum Organizations. . . . 26
Researchers. 30
B. Needs for Assessment Data at the State Level. 38
C. Needs for Assessment Data at the Local Level. 47

SECTION II Ways To Enhance the Utility of the National Assessment 55
Dissemination and Data Accessibility 63
Increased Interpretation and Data Analysis 64
Increased Background and Program Variables. 66
Increased Research Potential: Sample and Package Design. 67
Increased Policy Relevance 69
State and Local Needs. 70

SECTION III Summary 73

APPENDIX A Selected Examples of Uses That Have Been Made of National Assessment Data and Materials 75
National 75
Regional 80
State. 81
Local. 87

APPENDIX B People and/or Organizations Contacted by Abt Associates Inc. and by the Education Commission of the States 91
People Interviewed by Abt Associates Inc.. 91
People Interviewed by the Education Commission of the States 94

BIBLIOGRAPHY 97

LIST OF TABLES AND EXHIBITS

TABLE 1. Needs for, Utilization of and Suggestions for Enhancements of National Assessment: Federal Legislative Branch. 11

TABLE 2. Needs for, Utilization of and Suggestions for Enhancements of National Assessment: Federal Agencies. 16

TABLE 3. Needs for, Utilization of and Suggestions for Enhancements of National Assessment: National Education Organizations. 22

TABLE 4. Needs for, Utilization of and Suggestions for Enhancements of National Assessment: National Curriculum Organizations. 28

TABLE 5. Needs for, Utilization of and Suggestions for Enhancements of National Assessment: Education Researchers 32

TABLE 6. Needs for, Utilization of and Suggestions for Enhancements of National Assessment: State Level 40

TABLE 7. Needs for, Utilization of and Suggestions for Enhancements of National Assessment: Local Level 48

TABLE 8. Changes in the National Assessment Model and Potential Trade Offs 57

EXHIBIT 1. Ways of Providing Information. 62

PREFACE

The following paper is one of seven required by the National Institute of Education (NIE) grant to the Education Commission of the States for the administration of the National Assessment of Educational Progress (NAEP). The papers are intended to raise significant questions about NAEP functions and procedures in order to apprise the Assessment Policy Committee of its technical and administrative options as NAEP moves into a second decade of activity. According to the NIE Program Announcement, each paper should include "discussions about the significance of a national assessment of educational progress for (a) informing public opinion, (b) influencing educational and other social policies, (c) identifying specific educational and research needs and (d) contributing to the goal of improving education." Thus, the papers combine broad theoretical considerations with very concrete discussions of potential options and their trade offs. All the papers are suggestive, rather than definitive -- starting points for discussion rather than end points.

Two fundamental questions undergird these papers. First, if one were to design a national assessment program from scratch to serve the needs of the eighties and nineties, what would it look like? What theory would it draw upon? What constituents would it serve? How would it operate? And second, given that an assessment has been operating for 10 years, how do answers to these questions match the current NAEP? How could the current NAEP be changed to accommodate needs and technology unforseen when it was begun in the 1960s? As these questions weave in and out of the discussions of various issues, it will become clear that the answers one gives will probably depend upon one's assumptions about such things as the way society, education and science work, as well as one's familiarity with the theory, politics, history and operation of the current National Assessment.

There are many ways to implement the 1979 legislation mandating a national assessment. Some are undoubtedly better than others in certain respects; all have both advantages and drawbacks, complicating even the

NOV 17 1980

simplest comparisons among alternatives. In such a situation, it is imperative that advocates for various models of a national assessment adopt an open, interdisciplinary approach to rethinking the Assessment, listening carefully to each other and respecting professional differences of opinion.

Individuals charged with making policy decisions about NAEP -- members of the Assessment Policy Committee and members of the U.S. Congress -- will find that, in many instances, attempts to strengthen the Assessment in one regard will weaken it in some other regard. Given a certain level of funding, advances in some areas come at the expense of others. The National Assessment can neither be all things to all people nor a single, unchanging enterprise. It will always represent a combination of various tensions. The question for its future is not how do we eliminate those tensions? Rather it is: How do we harness those tensions creatively to keep the Assessment abreast of the times and continually relevant to the needs of American education?

INTRODUCTION

Education in the United States is a diverse enterprise. Responsibility for education is reserved to the states under the U.S. Constitution. Local school boards have the responsibility for making and implementing decisions for their districts; state agencies provide advice and guidance, administer statewide programs and control expenditures of state funds. Federal programs, such as Head Start and Title I and their associated guidelines, also have a major impact. Program emphases at the national level -- for example, career education -- are often reflected in some way in the local schools' activities. In addition, education is everyone's business. As parents, students, taxpayers or simply concerned citizens, nearly everyone has some interest in what the schools are doing and how public monies are being spent.

During the 1960s, people's perceptions of the type of data needed to make decisions about the schools shifted. Previous data collection efforts had primarily centered upon inputs to the education system -- buildings, teacher salaries, teacher/student ratios -- with the assumption that improved inputs would result in improved outcomes, i.e., better student achievement. As it became apparent that this assumption did not necessarily hold, attention was directed toward measuring educational outputs, or student achievement, as an indication of how well the schools were doing their job.

As a part of this movement, the National Assessment of Educational Progress (NAEP) was designed in the mid-sixties to provide nationwide data on student achievement and to monitor changes in achievement levels over time. The developers of the Assessment felt that nationwide data would be useful since states or school districts could then examine their own programs to determine whether the same trends were taking place. National policy makers could, it was thought, use nationwide data in program planning because they would have information about students' achievement.

As we enter the 1980s, it is becoming more and more apparent that no one type of information about the

schools is a sufficient basis for decisions. No one study, research effort or experiment can come close to providing the answer about what the schools should be doing; in fact, the complexity of the problem often seems to increase as we gather more data about it.

This paper examines needs for achievement data expressed by various groups involved with education and examines whether and how National Assessment might meet those needs. In examining the roles that National Assessment might fill, it is necessary to remember that the Assessment cannot and could not, even with unlimited resources, succeed in being everything to everyone and that choices will have to be made. These choices will be made by the Assessment Policy Committee in accord with the priorities that it establishes for the Assessment.

The National Assessment is designed to collect and report data on the progress of education, nationwide. To do this, the Assessment measures the abilities of 9-, 13- and 17-year-old students (and out-of-school 17-year-olds and adults aged 26-35 when resources permit) in learning areas traditionally taught in the schools. Learning areas currently assessed are: mathematics, writing, reading/literature, science, citizenship/social studies, art, music and career development.

In the 1978 Education Amendments, the U.S. Congress specified the National Assessment's primary responsibilities: (1) to collect and report data on achievement in reading, communication and mathematics; (2) to provide data on changes in achievement in these areas over time; (3) to conduct assessments in other areas as the need arises; and (4) to provide technical assistance to state and local education agencies in using Assessment objectives and making comparisons with national data.

The major function of the National Assessment has been the detection of trends in education achievement over time. This function has shaped many features of the Assessment design. The Assessment is a large-scale survey intended to provide information on what is happening in the schools. Since it is impossible to predict areas in which changes might occur, it is important that the Assessment take measurements in a number of subject areas and in a variety of topics within those subject areas so that any changes in achievement that do occur can be detected. The detection of small changes -- those as small as 1 or 2 % -- demands considerable measurement accuracy and precision. If the data are to

be relied upon, procedures must be duplicated as nearly as possible in each assessment so that changes are not attributable to changes in procedures. These considerations led the designers of the Assessment to create an assessment that is broad -- in the sense that it covers many subject areas and many topics within subject areas -- and relatively stable in terms of procedures and items.

Utility of the National Assessment

In considering the utility of National Assessment, we must consider some of the very real limitations of social science data and some of the research constraints imposed by the present design of the Assessment. The nature of social science data is such that "most policy-oriented research, at least in education, tends to influence the broad assumptions and beliefs underlying policies, not particular decisions" (Cohen and Garet, 1975, p. 39).

National Assessment was conceived during the early 1960s, a period when "there was much hoopla about the rationality that social science would bring to the untidy world of government" (Weiss, 1977, p. 4). However, as the 1960s wore into the 1970s, it became apparent that the impact of social science data had been, to say the least, less than was hoped for (Patton, 1978, pp. 18,19). Armor notes the inadequacy of social science survey data in establishing cause-and-effect relationships:

> The establishment of cause-and-effect relationships generally requires an experimental, longitudinal study design. Survey data can establish, with certainty, concomitant relationships at best. Any assertions of causal order are purely speculative ... (Armor, 1972, p. 170).

Although survey data may not be optimally effective in answering specific questions, this is not to dismiss their usefulness out of hand. Survey data of the type provided by National Assessment may be most effective not in providing information for specific decisions but for providing a general background or climate of information which, taken together with other sources of data, will lead to changes in public perceptions and eventually to changes in programs and policies. In fact, there are evidences that this type of influence has

taken place. The National Council of Teachers of Mathematics drew upon NAEP data, which supported data from other sources, in developing its curriculum for the 1980s. Mathematics teachers had indications that problem-solving was not receiving sufficient emphasis in the schools; NAEP data confirmed this belief. National Assessment writing data appeared at a time when national concern about writing achievement was growing; NAEP data and methodology, in conjunction with data from other sources, have influenced the development of ways to measure writing performance and have encouraged educators to pay more attention to the development of writing skills.

In considering research use, the strengths and limitations of survey data must be kept in mind. The Assessment is not designed to be an instrument for causal research or hypothesis testing. Assessment data provide descriptive information about conditions as they presently exist. These data cannot be used to attribute causation because factors are not controlled -- their occurrence is simply described.

Assessment data might best be regarded as a source of information for generating hypotheses rather than testing them. Assessment data can be used to describe relationships among various factors. Such information can be used in generating hypotheses regarding the relationships discovered; these hypotheses can then be tested using smaller-scale studies. Burton describes the difference between exploratory and confirmatory research:

> Confirmatory research is deductive; it goes from general principles to more and more complex specific applications. In exploratory research, one starts at the data and attempts to move toward generalization. One attempts, if possible, to move upward in generality by adding assumptions and discarding details -- but tentatively, since one has no accepted theory to separate the wheat from the chaff (Burton, 1978, p. 9).

There are many cases in which controlled experimentation is not possible. Broad social changes cannot be dealt with experimentally because they are not within the control of an experimenter. However, they may fall within the scope of survey data such as NAEP's. For example, the large-scale, court-ordered desegregation of schools in the South occurred between the first and sec-

ond national science assessments. The mass use of electronic calculators was preceded by the first mathematics assessment. Assessment data allow a unique perspective on the possible effects of such social changes.

Survey of Assessment Uses

In preparing this paper, the Education Commission of the States (ECS) sought opinions of a number of Assessment users and potential users to determine their needs for achievement data, their use of National Assessment and ways in which the Assessment might be enhanced to be more useful in meeting their needs. Audiences were identified at the national, state and local levels. The Education Commission contracted with Abt Associates Inc. of Washington, D.C., to interview users at the national level. Abt staff talked with members of the legislative branch at the federal level, federal agencies and national education organizations. The Education Commission contacted national curriculum organizations, education researchers and persons involved with education at the state and local levels. Some interviews were conducted in person; others were conducted by telephone. In addition to responses from persons interviewed, the paper draws on the experiences of National Assessment staff in working with various audiences and the comments, suggestions and recommendations that staff has gathered from these audiences over the years.

Persons and/or organizations contacted are listed in Appendix B. To encourage free exchanges of opinions, respondents were assured that they would not be directly quoted by name. It also should be noted that the views expressed are personal ones and do not necessarily reflect the position of the organization or agency with which a person is associated. Interviewees were told of the nature of the papers and were included in the survey only if they expressed willingness to participate. Nearly all of the persons contacted were willing to participate.

The persons and organizations contacted are but a small fraction of those who might have been included. The Education Commission chose to focus primarily on people who had had some experience with National Assessment, believing that in many cases people who were not familiar with National Assessment might have difficulty determining how it might be improved to meet their needs. Attention was devoted to audiences specifically involved in education and, for the most part, to groups

with a responsibility for or a strong influence in education decision making.

The following section describes different audiences' needs for National Assessment information, their uses of National Assessment and their suggestions for enhancing the utility of the program. Section II summarizes suggestions for potential enhancements to the Assessment program and examines the possible trade offs involved.

SECTION I

NEEDS FOR, UTILIZATION OF AND SUGGESTED ENHANCEMENTS TO THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

The general consensus found in responses by persons interviewed was that while National Assessment data are by no means the only source of data on education achievement, they are a unique source and fill an informational need not met by other sources. National Assessment provides information on the achievement of the entire national population at specific age levels in a variety of subject areas. It provides data on a much broader range of abilities than do norm-referenced, standardized tests, which typically emphasize verbal and mathematics abilities, not achievement, knowledge or skills in particular subject areas.

The question of need for National Assessment information is tied to that of utility of the information. Accordingly, National Assessment considered ways that the Assessment has been used in the past and is being currently used as one indication of needs. Appendix A gives a brief summary of ways that National Assessment has been used, focusing primarily on the last five years.

The following discussion describes utilization of and needs for National Assessment data and materials for audiences at the national, state and local levels and presents their suggestions for enhancing the Assessment in accord with their needs.

A. NEEDS FOR ASSESSMENT DATA AT THE NATIONAL LEVEL

Several distinct audiences with differing needs have been identified at the national level: legislators and legislative staffs, federal agencies, national education and curriculum organizations, and educational researchers. Abt Associates Inc. provided ECS with draft reports of their discussions with representatives of the federal legislative branch, federal agencies and national education organizations. These draft reports

have been summarized by ECS staff for inclusion in this paper. A complete report of these interviews, prepared by Abt Associates Inc., will be available later in 1980.

Federal Legislative Branch

Staff from Abt Associates Inc. talked with Congressional staff, personnel of the Congressional Research Service of the Library of Congress and personnel of the Congressional Budget Office. For the most part, Congressional staff interviewed are on the staffs of the relevant standing committees. In the U.S. House of Representatives, staff were from both the major committee responsible for authorizing education programs (Committee on Education and Labor) and the major appropriations subcommittee (Labor-HEW subcommittee of the Committee on Appropriations). In the U.S. Senate, staff were also from the major authorization committee (Labor and Human Resources) and appropriation subcommittee (Labor-HEW).

The Congressional Research Service of the Library of Congress and the Congressional Budget Office are service organizations providing information to the Congress in general and the committees in particular.

Interviews were conducted with 15 members of the legislative branch. Persons interviewed are listed in Appendix B. Table 1 summarizes their uses of the Assessment and suggestions for possible enhancements.

TABLE 1. Needs for, Utilization of and Suggestions for Enhancements of National Assessment: Federal Legislative Branch

Needs	Utilization	Enhancements
General information about student achievement and trends.	General background information -- testimony on national proficiency testing and mathematics results.	Briefer and more journalistic reporting.
Need for data to use as supporting evidence in speeches, presentations, decisions, etc.	Consideration of NAEP for use in Title I evaluation.	More interpretation of results.
Need change data much more than baseline data.	Cited NAEP data in speeches.	Coordinate timing of releases with federal legislative calendar; inform relevant people prior to releases.
Information on federal programs, bilingual students, high school facilities and programs; relationships with student achievement.	General background -- aware of NAEP through media.	Information on high schools.
		Information on bilingual students.
		Smaller geographic breakdowns -- 10 federal regions perhaps.
		Report data on public schools v. private schools.
		Information on school-related variables.
		Gather data on out-of-school youths and young adults.
		Gather information on whether students are receiving special services in the area in which they are tested (Title I, Follow-Through).
		Basic skills as primary Assessment focus; obtain funding for other areas from other sources.

Utilization of National Assessment by the Legislative Branch

Many respondents stated that they had not used NAEP data to date because the data are not specifically relevant to federal programs. One appropriations staff member, while making that comment, also said that his ability to use the data at the federal level was not important *if* the data were used by people at the state and local levels.

The National Assessment has been the subject of three specific uses by the legislative branch in the past. The first occurred in 1973 when Congressman Al Quie, then the ranking Republican on the House Committee on Education and Labor, proposed to change ESEA Title I to a program using education need as the major factor in funds distribution. He further suggested that NAEP was a model that could be used to allocate dollars among states. Because of that proposal, House staff became quite familiar with the workings of NAEP. The NAEP director was among the witnesses called to testify at that time.

In 1977, Senator Claiborne Pell proposed, in part based upon testimony from Admiral Hyman Rickover, that a national proficiency standard be established. The NAEP director was again called as a witness and NAEP data were cited during Senate hearings. Within the past year, Representative Carl Perkins, chairman of the House Elementary, Secondary and Vocational Education Subcommittee, held a hearing on the topic of math proficiency inspired by the recent release of a NAEP report. In the last two instances cited, NAEP data were used to provide general background information. No specific action was expected to be taken as the result of NAEP data.

A large number of respondents also mentioned their use of NAEP data for speeches. One said that the data are very useful "as a foil against SAT and minimal competency."

Virtually all of those interviewed received information on NAEP, most frequently by receiving the newsletter. Some staff have also requested additional information, but that has been an infrequent occurrence. Respondents were about equally divided between those who felt that briefings, during which Assessment results would be presented, would be useful and those who were content to continue to receive mailings. On balance, it would seem that offering a briefing in Washington on

Capitol Hill would be a service to a number of people. Every respondent also mentioned being aware of NAEP through press reports.

A number of respondents mentioned that NAEP data are especially useful because NAEP has the only data base that covers a truly random sample of the population. According to these respondents, especially those in the service agencies, this makes NAEP uniquely valuable in answering questions on such matters as test score decline.

Without exception, every respondent replied that NAEP change data are most valuable because the data permit comparisons. Several respondents mentioned that change data are virtually the only data to which they pay much attention. Several also said that more emphasis should be placed on change data in the basic skill areas and that other subject areas should be reduced or eliminated unless paid for by others (U.S. Department of Energy or Labor, or the National Endowment for the Arts, etc).

The frequency of NAEP's use by Congressional staff members can best be described as sporadic. In virtually every case in which NAEP was used, it was tied either to the release of data or to the emergence of specific issues through the legislative process.

Suggestions for Enhancing National Assessment in Accord With Needs of the Legislative Branch

Several respondents felt that the reporting of NAEP data was often too technical and should be made more journalistic. It was also suggested that NAEP should take a much greater initiative in interpreting the data. Respondents mentioned the fact that the timing of release of NAEP data should be better coordinated with the federal legislative calendar; three respondents also suggested that it would be most helpful if NAEP informed relevant people at the federal level of the planned release of NAEP data at least six months in advance of the release.

It was suggested that ECS make available literature on the strengths and weaknesses of the test. It was noted that the Educational Testing Service already makes this kind of information available. About one-quarter of the legislative branch respondents suggested that separate information be made available on students who

are bilingual. Another respondent also suggested that NAEP gather information on the high schools from which tested students came. Other suggestions were to report data on public school compared to private school students, to gather data on hours of instructional time per curriculum area and to obtain data on districts undergoing court-ordered desegregation.

The greatest number of suggestions involved changes that would impact policy and the allocation of resources. Some respondents suggested that despite the agreements that state-level data not be gathered, such data would be most interesting. However, these same respondents also said that the collection and release of such data would be very controversial and, in almost the same breath, dismissed the idea. One respondent suggested that NAEP report the state-level data that it does have to the states and allow the states to use the data in whatever way they thought appropriate.

It was felt that the regions used by NAEP are too large. One interesting comparison might be to report data by the 10 federal regions rather than by the four census areas. A number of respondents suggested that even if state-level data were not used the sample size should be increased substantially to allow for greater accuracy when finer breakdowns are done.

About one-third of the legislative branch respondents said that more needs to be done to obtain data on out-of-school youths and young adults. This issue was specifically mentioned in connection with the current debate on the youth initiative. Staff felt that they do not have any real information about the skill levels and education needs of those no longer in the education system. A related need is for data on adult illiteracy. One key staff respondent felt that the young adult assessment previously conducted by NAEP could have provided some very valuable information on that issue.

Most felt that the basic skill areas (reading, math, writing) should be the primary focus of the Assessment. Several suggested that knowledge about other nations also should be a priority. Four respondents in the legislative sample would like to know whether or not tested children are receiving any special services in the area in which they are tested (Title I, Follow-Through, etc.). These respondents, for the most part, felt that such data would give them valuable information on the impact of federal programs over time. Another,

and somewhat related, suggestion is that the same students be kept in the sample over time so that a data base on the impacts of schooling and special services intervention could be obtained.

Seven of those interviewed said that NAEP is of value. A typical comment was, "If we didn't have it, it would have to be invented." Another respondent said that it was doubtful if a policy decision had ever been made on the basis of NAEP data alone but that the National Assessment does serve a valuable catalytic function, stating in addition that the most important outcomes of NAEP are probably the debates that test results trigger. Several of the respondents felt that NAEP is too expensive for the quantity and quality of data obtained.

While some respondents did feel that ECS is an appropriate home for the Assessment, others questioned the ECS commitment and felt that ECS had no interest in the evolution of the program, noting that NAEP is virtually an autonomous unit within ECS. The Carnegie Corporation and the University of Iowa were two potential sponsors mentioned.

Federal Agencies

Twelve representatives from federal agencies were interviewed, representing offices responsible for research and evaluation, policy and program administration. Persons surveyed are listed in Appendix B. All were familiar with the National Assessment project and had used NAEP data. Table 2 summarizes needs, utilization and suggested enhancements for the National Assessment expressed by these representatives.

TABLE 2. Needs for, Utilization of and Suggestions for Enhancements of National Assessment: Federal Agencies

Needs	Utilization	Enhancements
Needs for general background information on achievement and achievement trends.	General background information.	Addition of new assessment areas -- e.g., speaking and listening.
Information to use in support of specific programs and policies.	Used in preparation of NIE research program on reading.	Addition of new populations -- e.g., adults, Hispanos, handicapped.
Data to use in evaluating programs and policies.	Science measures used by NSF.	Enlarge samples and collect more background information -- report cross-tabular information.
Need for information on thinking and problem-solving.	Cited in speeches and papers.	Sample smaller identifiable populations -- e.g., states, bilingual, Title I programs, basic skills programs.
Need for change data and data on basic skills.	One agency commissioned Assessment and uses results.	Increase amount of contact between NAEP, federal agencies and researchers.
	Used in compensatory education study; sustaining effects study; in formulating new youth bill.	Increase possibility of doing studies in conjunction with governmental agencies.
	Used to formulate an administration position on test score decline and federal perspective on testing.	Increase research utilization of Assessment.
	Used in annual evaluation reports and Condition of Education.	
	Used to prepare testimony, identify program needs.	

Utilization of National Assessment by Federal Agencies

Most respondents saw the principal value of NAEP as a source of general background information. It is the most comprehensive measure of education output available and thus provides a context within which specific policies can be considered.

Some specific applications were also mentioned. Several agencies have contracted with the National Assessment to include items in an assessment -- for example, the Food and Drug Administration administered several items on the ability to read labels. The National Institute of Education used NAEP materials in the preparation of its research program in reading; the National Science Foundation has used the science measures. Data were used in a compensatory education study and a sustaining effects study, in evaluating the meaning of test score declines, in preparing testimony and in formulating bills. National Assessment data also were used in preparing an administration position on test score declines and the federal perspective on testing. Assessment data are cited in The Condition of Education, in speeches and papers and in annual evaluation reports.

Although they did cite some specific uses, agency representatives were more inclined to view NAEP as a source of background information. While the NAEP data by themselves are not the only or even the main determinant in the development of policy, they are a resource that is widely used within the federal government to address policy positions. Nearly all the respondents use NAEP reports as a resource in writing speeches and memos.

Most respondents surveyed identified the assessments of basic skills -- reading, writing, mathematics and science -- as most important to them. Agencies concerned with consumer education, career education and basic skills were interested in data in those areas. Although respondents were not against the assessment of other areas, they felt that areas such as art and music could be dropped, if budget reductions necessitated such cutbacks.

Federal agencies view the reporting of changes over time as the major value of the National Assessment. All respondents cited the value of the change data. The breakdowns that were of most interest for these respondents were age (especially the 17-year-old group), race,

region, community size and sex. Some respondents felt that closer coordination of NAEP reports with legislative cycles might improve their usefulness; others felt that most changes occurred slowly and that measurement every five years was sufficient.

Suggestions for Enhancing National Assessment in Accord With Needs of Federal Agencies

The major theme identified by all the respondents is the desire to have the data made more useful in an analytical sense. Recommendations ranged from suggesting more interpretation of results at press conferences to a total restructuring of the project.

Most felt that NAEP's general communication efforts are good. All were familiar with newspaper reports on NAEP results; some had used reports and had personal communications with NAEP staff, including briefing sessions. The sessions for Congressional staff and press conferences were identified as especially successful.

Respondents suggested that communication between NAEP, federal policy makers and researchers should be improved. The national education policy audience and the National Assessment need to be in closer touch in order to anticipate areas of concern and investigate how NAEP might be of assistance in those areas. One respondent reported that his agency had always meant to get together with NAEP to design some mutual studies but had just not gotten to it. Two respondents said that a quick-access data base would be most helpful for the short-term studies that one often needs when dealing with policy concerns.

One respondent made the point that since NAEP data are of general rather than specific usefulness, NAEP will have to be particularly aggressive in pointing out the value of the data to federal agencies.

Another respondent suggested that NAEP should prepare a report to each federal agency having interests in the areas covered by a particular assessment, and NAEP staff should meet with staff at the level of the Secretary to discuss the reports in depth. Issues to be discussed with key departmental personnel might include policy issues that need to be addressed in the next few years, data that are needed to support these policy issues and the data that NAEP could collect to support such issues.

Several respondents thought that National Assessment is too sensitive in considering federal requests for data or information. For example, a federal agency wanted to include questions about international education in an assessment and was turned down. Some agency representatives felt that federal requests for information did not seem to convey the same legitimacy as requests from other audiences.

Survey respondents felt that use of NAEP data by outside researchers should be encouraged. Difficulties in using NAEP data tapes were cited as one possible cause for the lack of use. Work is presently underway to improve data tapes. There was some feeling that ECS should not be responsible for secondary data analysis but should provide access to the data to outsiders. It was also suggested that NAEP items could be used in other research studies. For example, the 1980 Longitudinal Study contains some NAEP items.

Other suggestions had to do with potential additions or changes to the NAEP program. One respondent suggested adding new assessment areas -- for example, oral evaluation. Another suggestion was to add new populations -- for example, the handicapped -- to retain the adult sample and to enlarge the sample of Hispanos to allow more complete data reporting. Utility might also be enhanced if NAEP could enlarge its samples and collect more background data on respondents. Cross-tabular information was cited as a type of data that would be particularly useful (urban blacks v. rural blacks; Southeastern males v. Southeastern females, etc.). Although some cross-tabular information can now be obtained using National Assessment data, it has not been reported in many instances and, in some cases, is only available for summaries over groups of exercises, not for individual exercises.

Another set of suggestions had to do with providing information on smaller identifiable populations. Suggested populations were: states; participants in federally supported programs, e.g., bilingual programs, Title I, basic skills; school districts. A number of respondents reported that the lack of state-level data restricted the usefulness of the Assessment. They acknowledged that most states are resistant to having these data reported, but the respondents nevertheless believe that the data would ultimately benefit education.

Other suggestions had to do with changing the way children are assessed. These suggestions concerned the collection of data on the same individuals over time. Longitudinal studies would then be possible; information on schools, parents and personal characteristics could be related to achievement. The respondents who suggested this recognized that there would be large costs and that NAEP would essentially be a different project. It was also suggested that the collection of NAEP data be structured around hypothesis testing and related to a theory. This might make data more interesting and valuable; however, it was noted that this would also dramatically affect the present design of the project.

All of the suggestions were identified as ways to make NAEP more useful to policy makers, program managers and researchers at the federal level. Respondents realized that many of their recommendations would be difficult to achieve, and most felt that nothing should be changed that would obscure or be detrimental to NAEP's value as a general measure of educational progress.

All the respondents indicated that NAEP is important and should be continued. Several felt that the cost effectiveness of the effort relative to other efforts should be monitored. There was a great deal of hope that Assessment results could become more useful to those trying to solve problems in education at the federal level.

National Education Organizations

The national education organizations surveyed by Abt Associates Inc. ranged from those involved with curriculum development to those concerned with educational research to those concerned with the interests of elementary and secondary teachers. A list of the organizations surveyed is included in Appendix B. Some organizations act largely as lobbyists on Capitol Hill; others serve mainly a liaison function between their membership and various federal agencies; still others are primarily engaged in information exchange. The extent to which NAEP data are used depends upon the organization's activities; those involved in all three of the above activities tended to use NAEP data more and to be more familiar with the design of the Assessment program than those involved in only one of the activities.

Among the 15 representatives of education organizations interviewed, those using Assessment data exten-

sively were in the minority. Use of data tended to be relatively infrequent, and generally was tied to NAEP releases or need for data on specific issues. Table 3 summarizes needs for, utilization of and suggested enhancements to the National Assessment expressed by national education organizations.

TABLE 3. Needs for, Utilization of and Suggestions for Enhancements of National Assessment: National Education Organizations

Needs	Utilization	Enhancements
General background information on status of education.	Background information in speeches, papers, etc.	Include executive summaries in reports.
Information on trends in education achievement.	Findings reported to membership via organizational newsletters.	Improve dissemination -- present more analysis in news releases to prevent distortions by news media.
Information on issues of importance to the particular membership of the organization.	Used indirectly as an influence on policy.	More interpretation of results -- help state and local practitioners use data.
Information useful in making recommendations to state and local practitioners.	Sometimes used in formulating positions, priorities, recommended policies.	Provide and disseminate more information about the design of the project.
		Provide direct assistance to local users -- perhaps in conjunction with national education organizations.
		Provide additional breakdowns of categories -- state-by-state data, public v. private schools, family income, type of school district (rural, urban, suburban).

Utilization of National Assessment by National Education Organizations

Although the associations do not make extensive use of NAEP, everyone surveyed was familiar with NAEP, had a general knowledge of the type of data collected and valued the Assessment program for providing them with a sense of education achievement in the country. The most common use of NAEP data mentioned was the reporting of the results of the assessments in an organization's newsletter or magazine. Generally, the organizations select those data that pertain most to the focus of the association -- e.g., achievement data for elementary school children or achievement data for the basic skills curriculum areas -- in framing reports to their memberships. The associations, without exception, use the trend data more than the data from an individual assessment. They usually do not write an analysis to accompany the data, but publish the data in a reportorial fashion.

Ten of those interviewed said that the data act as a general indicator of the state of education nationally. Although the data are not often used as the basis for policy positions, those interviewed said that NAEP influences their thinking about education and thus indirectly affects the positions they take on certain issues. Two individuals did mention they used the data in forming policy, either for their organization internally or in policies they were advocating on Capitol Hill, although they said this type of use is infrequent.

Four staff members said that they use the data as background data in papers or in speeches to constituents. The data are not used to indicate the value or importance of particular positions, but are used as a backdrop against which other points are made or issues explored. The data are not thought by themselves to be persuasive, but are thought to be useful in building an argument or setting the foundation for arguments that are meant to persuade.

The associations do not frequently use other sources of achievement data either. In fact, in all but one case, associations' use of other test data was always mentioned as less frequent than their use of NAEP data. Other data used were the Educational Testing Service test data, the Iowa Test of Basic Skills and the California Achievement Test. Organizations felt that NAEP data were more appropriate to their needs than data from these other achievement tests for a number of reasons.

First, NAEP data cannot be used to measure teacher effectiveness because the data are national and the Assessment is objective-referenced rather than norm-referenced. Second, NAEP presents trend data, which are thought to be more realistic measures of changes in student achievement than change measures that are sometimes obtained using data from the other tests. Third, some respondents valued the ability to cite achievement data for particular groups of students, either by age, race or geographic region. These breakdowns, often not available for the other tests, may make the data more meaningful to the memberships of these organizations. Finally, two education association staff members noted that NAEP assesses students' applied skills more than do the norm-referenced, standardized tests.

Suggestions for Enhancing National Assessment in Accord With Needs of National Education Organizations

Since many organizations did not make extensive use of NAEP data, they did not have many suggestions for changing the design of the Assessment. Most comments centered on reporting and disseminating data.

About half of the respondents are on the NAEP mailing list and thus receive the NAEP Newsletter. Most expressed interest in receiving the Newsletter, saying that it highlights the results of each assessment and lists the reports available, enabling them to order reports if more detailed information were needed.

Several respondents suggested possible improvements in NAEP's dissemination plan. Three respondents thought that the media reports on the results of the National Assessment often distorted the data. Although it was acknowledged that NAEP staff ultimately have little control over what is reported in the press, respondents suggested that NAEP staff try to write future press releases of sufficient depth and in a style that would encourage use of the releases as originally written.

Two individuals also indicated a need for more summary information. It was suggested that reports might contain an executive summary, thus enabling those who either did not have the time or the interest in reading the entire report to get an overview of the Assessment results.

Some respondents thought that NAEP should provide more interpretive analysis of the Assessment data, either in their reports or through technical assistance or briefings. State and local educational practitioners, who make up the the majority of the membership of these national organizations, often are uncertain about the meaning that national-level data have for them. The many questions that respondents asked about the design of the program and about the representativeness of the sample also indicate that the associations need more information on the NAEP design.

Eight of the respondents felt that the NAEP staff should direct some of its reports to specific audiences, particularly education practitioners at the state and local levels. Two respondents said that the data ought to be presented in a form more usable to education researchers. Coupled with this, some respondents also mentioned that NAEP staff might provide direct technical assistance or briefings to local audiences on the potential uses of Assessment data. These efforts to reach local audiences could be facilitated by working with national education associations.

Several respondents mentioned that additional reporting categories would be useful. Two respondents said that state-level data would be valuable for their constituents. They noted that although such data could be harmful if used politically, the data might be beneficial to state education practitioners, who could then use the data in setting state priorities. Three respondents noted that disaggregating the data to show education achievement for private school children separate from public school children would be useful. Without this breakdown, these respondents said, the data are not very meaningful for private school educators because the curriculum emphasis in private schools is different from that in public schools. One respondent also mentioned that breaking the data down by type of school district -- e.g., urban, suburban or rural -- and by family income level would be useful.

The interviews with staff of national education associations clearly indicate that the NAEP program is generally respected by the national education associations. The limited use made of the data, moreover, does not reflect the value that these associations place on the Assessment. Those spoken with believe NAEP to be more meaningful for national policy makers than for state and local practitioners. Thus, the associations are reluctant to, and feel unable to, use the data when

advocating policy positions that are meant to serve their membership groups, which are in large part comprised of practitioners at the state and local levels. The respondents, however, did note that the Assessment might have the potential for being valuable to practitioners at the state and local levels and suggested that NAEP staff make a more concerted effort, through the national associations, to reach these individuals.

National Curriculum Organizations

In considering needs of national education organizations, we have made a distinction between organizations with a general concern for education whose members are part of a specific constituency -- for example, the National Education Association, the National Association of State Boards of Education and the Association for Supervision and Curriculum Development -- and organizations with responsibilities connected with a particular subject area, since needs for Assessment data and utilization of the Assessment vary somewhat between these two groups.

National Assessment staff have had extensive contacts with members of many national curriculum organizations during the development of assessments in particular subject areas. In some cases, representatives of these organizations were not included in the current survey since information on their needs and suggestions for enhancements had been obtained in earlier contacts.

Utilization of National Assessment by National Curriculum Organizations

Some of the organizations that NAEP has worked with closely are: the National Council of Teachers of Mathematics (NCTM), the National Council of Teachers of English, the National Association of Science Teachers, Music Educators National Conference, International Reading Association and the National Council for the Social Studies. Table 4 summarizes needs for, utilization of and suggested enhancements to the assessment program expressed by national curriculum organizations. Currently, the primary uses of Assessment data by curriculum organizations are citing results in making recommendations regarding their areas and informing their memberships about Assessment results. For example, NAEP findings were used in support of the recommendations presented by NCTM in its curriculum for the 1980s; the

current emphasis on writing instruction was supported by NAEP data indicating a decline in writing skills. Assessment findings have been disseminated in the journals of many professional organizations: NCTM has presented a series of articles in the Arithmetic Teacher and the Mathematics Teacher, and articles on science findings have appeared in Science. The National Council of Teachers of English and the National Council for the Social Studies have also disseminated findings to their members.

TABLE 4. Needs for, Utilization of and Suggestions for Enhancements of National Assessment: National Curriculum Organizations

Needs	Utilization	Enhancements
To know how students are doing in many aspects of a subject area and to know about trends in the various aspects.	Cite findings, disseminate findings to membership.	Continued broad measurement across discrete subject areas. Expand capacity to permit users to reaggregate data.
To compare achievement in higher- v. lower-level cognitive processes.	Use findings in journal articles.	Continued broad measurement of change.
To examine the relationship of instructional variables and achievement; to obtain data on instructional variables.	Use findings as support for recommendations.	Increase collection of information on instructional variables.
To explore new methods of measurement in a subject area.	Disseminate methodology to those interested.	Either analyze or help others analyze relationships of instructional variables and achievement.
To find out impact of innovative programs.	Use objectives as basis for their own objectives; use to make recommendations on curricula.	Continue use of open-ended items when appropriate so diagnostic information will be available.
To gain information on the process of learning.		
To gain diagnostic information.		
To obtain objectives and items.		

Suggestions for Enhancing National Assessment in Accord With Needs of National Curriculum Organizations

Curriculum organizations see a need for a broad base of information about their particular subject area. A major asset of the National Assessment, in the view of subject-matter people, is that it includes both lower- and higher-level cognitive skills and is not restricted to narrow measures of ability. Curriculum organization members also felt that the inclusion of an area in the Assessment served to reinforce its importance in the schools. Educators are concerned about the trend toward narrowing education to the basics of reading, writing and arithmetic, and emphasized that the inclusion of many subject areas, broadly defined, provides vital support for the theory that education is more than just the basics.

Members of curriculum organizations also stated that information on changes in achievement, to be useful or even believable, must be based upon a sufficiently representative sample of items to accurately reflect the extent of the subject area. Changes in mathematics based only on computational skills or reports of social studies changes based only on knowledge of government would not be acceptable.

Many organizations expressed a need for more program information than is currently available. Most were more concerned with instructional and program variables than with student background variables. Information on students' course backgrounds, amount of instructional time in an area, instructional variables, school facilities and participation in specialized programs, such as writing labs, and how these relate to achievement would be useful. The curriculum organizations advocated obtaining as much information on instruction at the student and school levels as is consistent with maintaining broad measures of achievement.

National Assessment presently collects some information on instructional variables and is expanding this collection in future assessments; however, even when this information has been collected, it often has not been fully analyzed and reported. Respondents stressed that information on implementation of programs and on course-taking would be useful even if not related to achievement, since reliable, nationwide data on such factors often are not readily available.

Not all respondents agreed that additional program information would be an unmitigated good. One noted that "there's a danger in trying to tie results with programs. There are so many confounding variables."

National Assessment was encouraged to continue its exploration of innovative methods of measurement in the different subject areas. National Assessment has had the freedom to experiment with methods of assessing such things as music performance, essay writing, art production and attitudes, and the development of such technology has encouraged others to attempt measurement in these areas. As one example, the National Council of Teachers of English was instrumental in helping NAEP develop a system of scoring essays and has been very active in disseminating information about the system to interested users. However, respondents from curriculum organizations cautioned that NAEP should attempt to continue its development and use of innovative methods of measurement only insofar as that goal is compatible with retaining a broadly-based assessment covering many facets of a subject area.

Mathematics educators expressed the need for more open-ended items to aid in diagnosing errors students commonly make. Assessment of many aspects of writing also demands the use of open-ended items.

Members of curriculum organizations felt that NAEP provides a valuable resource. As one said, "It is a key, most dependable source for a truly national sample on basic literacy data." Another commented, "The opportunity to monitor change is a unique contribution of the Assessment. Without the Assessment we would have no continuing data base against which to examine curriculum and make adjustments to reflect priorities." Others emphasized that while data on mathematics and reading are readily available, the Assessment is the only nationwide source of data in many subject areas.

Researchers

The potential of National Assessment for research use is only beginning to be explored. Interpretive studies drawing upon NAEP data have been completed, and a number of researchers have used Assessment data for secondary analyses; however, efforts have been fragmented and a body of research using NAEP data has not yet

been generated. In addition, research efforts that have been completed generally have not been widely disseminated.

Education researchers work at the national, state and local levels. We have chosen to consider them primarily at the national level since National Assesment's data base is most suited for national-level research and since many researchers do work from a "national" perspective.

Presently, a team of researchers headed by a group from the University of Illinois at Chicago Circle are working with NAEP staff under a grant from the National Science Foundation (NSF) to explore the feasibility of secondary analyses of NAEP data. In conjunction with this grant, data tapes and codebooks are being developed that will standardize documentation and tapes and bring them more into conformance with other systems. In addition, the National Institute of Education (NIE) and the Education Commission of the States (ECS) will jointly award contracts to researchers during the summer of 1980 to do work using NAEP data. As the NIE and NSF projects are completed, much more information will be available on ways to enhance National Assessment's utility to researchers.

Several different types of research that can be accomplished using Assessment results and data have been identified -- interpretive studies, policy analysis and secondary analysis. Interpretive studies involve reviewing and manipulating existing Assessment results and data analyses to interpret results and put them in context. Policy analysis and secondary analysis both would generally involve additional data analysis. Table 5 summarizes the need for, utilization of and suggestions for enhancements of the Assessment for these three types of research.

In considering research uses of the Assessment, we have limited the discussion to uses that are possible using the present basic design of the National Assessment. Research needs that imply basic changes in NAEP will be addressed in Section II.

TABLE 5. Needs for, Utilization of and Suggestions for Enhancements of National Assessment: Education Researchers

Needs	Utilization	Enhancements
INTERPRETIVE RESEARCH		
Interpreting NAEP results to make them more useful.	Dissemination through professional journals.	Broader coverage of topics to allow firmer generalizations.
Linking performance with background and instructional factors.	Publication of reports designed to help people understand Assessment results.	Researcher input into development -- so interpretable groups of items would exist.
Linking Assessment results to other data.	Provide diagnostic information in some cases (NCTM journal articles).	More instructional variables.
Criteria for evaluating Assessment results.	Recommendations on practices (NCTM and Project Synthesis Studies).	Clearer presentation of items and data; clear lists of items.
POLICY ANALYSIS	Provide information for decision makers.	Publicize availability of data.
Information on effectiveness of programs.	Contribute to general climate of opinion about topical issues.	More open-ended items and nested items designed to be diagnostic.
Relationship of external, school and background factors to achievement.	Relate to other studies to draw and support conclusions.	Capacity to input to the Assessment on a relatively short-term basis.
Information on topical policy subjects -- nutrition, energy, etc.	Develop relationships and generate hypotheses using accepted statistical techniques.	Capacity to include policy questions in Assessment.
Information relating acquisition of knowledge to developmental and learning theory.	Add to base of information for use in policy decisions.	Background and demographic variables more closely tied to those of other studies -- income data, census region data, state data, grade data.
Information relating achievement and motivation.		Identifiers to allow linkages with other data sets (decennial census, CPS, ELSEGIS, OCR, Title I).
Information on generalizability of basic skills to subsequent learning.		Additional background variables -- ability to input to decisions on background variables.
SECONDARY ANALYSIS		Better access to data; better documentation of data and procedures.
Information related to a researcher's particular area of concern.		Increased publicity about availability of data.
Carefully drawn national sample, trend data.		Increased background variables, better SES indicators.
Access to student-level data.		Increased collection of school-level variables.
Guidelines for feasible and unfeasible studies using NAEP.		Opportunity for researchers to have input into Assessment and packaging design.
Easy access to data base.		

HOA

TABLE 5 -- *Continued*

Needs	Utilization	Enhancements
More background and instructional variables.		Increase size of sample -- finer demographic breakdowns, more information on social/ethnic groups.
More school-level data.		Collect income data and other background information to make compatible with other studies.
Increased comparability of variables with other studies (e.g., census definitions of urban/rural and community size).		
Increased attention to packaging considerations in line with research needs.		
Smaller demographic breakdowns (perhaps by state).		

Utilization of National Assessment by Researchers

Two recent studies point to the potential of National Assessment for use in conducting interpretive research. Project Synthesis, a recently completed project funded by the National Science Foundation to investigate needed new directions for science education, used NAEP data as an achievement indicator in conjunction with other data such as course enrollments and changes in science curricula to examine trends in science education and to make recommendations regarding changes in science education. A panel from the National Council of Teachers of Mathematics (NCTM) is currently interpreting results from the 1977-78 mathematics assessment (including some areas, such as hand-held calculators and attitudes toward mathematics, that were not reported upon by National Assessment) and will present their conclusions to the NCTM membership through NCTM journals.

National Assessment has not been widely used for direct policy relevant analyses. Some policy applications have been made using existing data; the potential of National Assessment for direct policy relevant analyses requires further investigation.

A number of researchers have completed secondary analyses of National Assessment data. The researchers who have worked with NAEP data have been eclectic in their interests: studies have ranged from statistical balancing of black/white data, to studying differential predictors of mathematics achievement for boys and girls, to examining students' political socialization and occupational socialization. The National Science Foundation studies will involve concerns in mathematics, science and the social sciences; one team is studying the effects of schooling on mathematics achievement (Welch et al, 1980).

The researchers contacted in the preparation of this paper were all people who had used NAEP data in some way and thus recognized and were prepared to deal with the limitations inherent in survey data. Most, therefore, addressed the question of improving the Assessment in accordance with its present purposes.

Suggestions for Enhancing National Assessment in Accord With Needs of Researchers

Most of the researchers surveyed expressed the opinion that National Assessment is a valuable data resource. One noted that "without the Assessment we would lose a national perspective and the long-term perspective that you don't get from standardized tests." However, while noting that the Assessment serves a useful function in giving a periodic and continuing review of education directions in the nation, one researcher stated that "its value would be increased if it could be put together with other pieces of information. Presently this can only be done intuitively."

Another researcher noted that for purely research purposes the Assessment is perhaps not the best vehicle. For the same expenditure of funds, numerous experimental studies could be conducted. However, the loss of the data for other audiences would probably offset any gains made from funding additional research studies.

Researchers stated that NAEP is well-suited to interpretive research and that the potential of this type of research has not been fully explored. As one said, "Much more could be done with the data as they already are. People tend to get tangled up in numbers." He added, "You can't use a single source of data. NAEP data used in conjunction with other sources represents a dimension not found elsewhere. No other data show how students are doing in specific learning areas within a discipline."

As Table 5 reveals, increase and improvement of background and instructional variables is needed. One researcher commented, "To increase research use, more or different school variables to relate to outcome variables are needed." Another noted, "Income data would make it much more compatible with the parameters of policy studies." Yet another termed the fact that there were no measures of parental occupation "scandalous."

Background variables should be made more compatible with other data bases. Some possibilities are: using census definitions of rural/urban and community size; relating NAEP demographic information to census definitions; obtaining parental occupational status or income information; and including identifiers to allow linkages with other data bases. Some researchers suggested that grade-level data would be more useful than age-level data, but there was no clear consensus on this point.

In addition to student background information, most researchers felt strongly that course background and program information should be collected, both at the student and the school level. As one researcher said, "There needs to be more attention to things that are manipulative so associations can be made." Some suggestions were: amount of homework, extent of TV watching, impact of length of "time on task," length of school year, relationship of students' labor force participation to achievement, and relationship of instructional practices to achievement. National Assessment has collected some data on TV watching and homework.

Needs for a broad base of achievement data and increased background variables may come into conflict. While those involved in curricular research need a broad coverage of the subject area, persons interested in relationships between achievement and background factors might be satisfied with fewer achievement indicators and more background variables. Several suggested that it might be beneficial in some of the Assessment item booklets (packages) to cut back on the number of achievement items. One said, "Typically a package is 90% achievement items. You might cut back to 65% achievement and 35% background items in some packages."

Most researchers stressed the need for finer geographical breakdowns. Noting that no one makes decisions at the regional levels defined by National Assessment, one researcher said, "The sample is not representative of policy units with which a policy analyst deals, such as states or school districts. A representative state sample would make the data much more useful." Others agreed that state-by-state breakdowns would be helpful, but stated that, failing this, smaller regional breakdowns would be an improvement.

Researchers also mentioned packaging and sampling considerations as drawbacks in using NAEP data. As one said, "No single package has enough measures on a single dimension to feel comfortable with." Anderson (1980, pp. 34, 35) notes the difficulty of constructing subtests using assessment packages. However, National Assessment purposefully distributes items measuring the same concept across packages to obtain more precise summary measures for these concepts. Another commented, "The sampling procedures, while extremely cost-efficient, make the data hard to use with the usual statistics. The jackknife procedure serves NAEP's purposes. How-

ever, the data may or may not be the best data for secondary analysis since the usual procedures don't apply."

Another common concern was that of accessibility to NAEP data. As mentioned previously, National Assessment is working with researchers at University of Illinois at Chicago Circle, under a grant from NSF, to standardize data tapes, code books and documentation. Those who had worked with previous data tapes noted that they were not easy to work with and that documentation was not always adequate.

All of the researchers contacted noted the cooperativeness of NAEP staff and complimented their willingness to give assistance. Most noted that extensive individual contacts and assistance are necessary to get the job done. One researcher commented, "Internal policies on dealing with researchers weren't all that clear. The organization wasn't geared up to do that well. The lack of policies is a problem."

Researchers also saw a need for a greater dissemination effort targeted specifically toward the research community. One respondent stated, "Dissemination is designed for public and informational consumption. It is not matched by the same level of communication with researchers. If researchers want information they generally have to go to NAEP." One suggestion was a mailing once or twice a year that described in technical detail changes in the data base, how the data has been used and so forth. Another suggestion was to provide prepublication announcements to researchers so that they would be aware of what was going to be available before reports appeared in the press. Researchers often are not aware of what is available.

Some types of research might demand more extensive changes in the Assessment. For example, those interested in research on a particular racial or ethnic group (American Indians or Orientals, for instance) saw a need for a larger sample of those groups than is presently surveyed; collecting information on program variables might also require a larger sample to insure that students who received the treatment were sampled. Fine-cell analysis (analysis of groups defined by more than one variable, e.g., sex X region, region X race) might also demand a larger sample. If only comparisons of fairly large groups are desired, a smaller sample than that presently employed might be feasible. Longitudinal data, classroom observations and individual interviews were also mentioned as possible enhancements to the As-

sessment.

Researchers also indicated that National Assessment itself should do more data analysis. One commented, "National Assessment has never figured out what could be done with the data to spark interest among those who might use it. It's kind of been written off." Another said, "You may be so caught up in building and processing the data that dissemination gets lost and analyses are not as creative as they might be." Another suggested that there perhaps is not enough interaction with other organizations, especially in the area of analysis. As one said, "Outside consultation is needed for background variables to gain insight into what has been done elsewhere. As much effort is needed on background variables as on item development."

B. NEEDS FOR ASSESSMENT DATA AT THE STATE LEVEL

The demand for education "accountability," which became increasingly evident in the 1970s, resulted in many states embarking on assessment programs to find out what their students actually were learning. Over 40 states presently conduct some type of state-level student testing program. The needs for these programs and the uses of the data obtained vary considerably from state to state, and state programs reflect this diversity. Some states test every student; some sample students but provide the opportunity for local agencies to "piggyback" with state assessments; some test every student in some areas and sample in others. In some states, the data are used primarily for program improvements; in others, school and district comparisons are made.

Usage of National Assessment information reflects the differing informational needs of the various states. Some states (for example, Minnesota) make extensive use of NAEP items to obtain comparisons to national and regional data. Other states (for example, Illinois) use some NAEP items in their assessment programs. Still other states review NAEP items and procedures and use or adapt them when appropriate.

Utilization of National Assessment at the State Level

Some of the states that presently are conducting or have conducted Assessment replications are: Arkansas,

Connecticut, Delaware, Florida, Hawaii, Illinois, Louisiana, Maine, Massachusetts, Minnesota, North Dakota, Rhode Island, Texas. Other states have used or adapted NAEP items for use in their assessment programs; many states drew upon NAEP methodology and procedures in designing their own assessment programs. Some of the state uses of National Assessment are summarized in Appendix A.

In preparing this paper, ECS limited its interviews to people in a state who were most likely to know about and to use Assessment data -- people concerned with state assessment programs. ECS also contacted the National Conference of State Legislatures to gain an idea of how state legislators use Assessment data.

Table 6 summarizes state-level needs for, utilization of and suggestions for enhancement of Assessment data. Some of the information in Table 6 was derived from experiences of NAEP staff in working with state personnel as well as from comments by survey participants. Seven state assessment officials were interviewed.

TABLE 6. Needs for, Utilization of and Suggestions for Enhancements of National Assessment: State Level

Needs	Utilization	Enhancements
Compare state's achievement to nation.	Comparison of state results to nation.	Finer geographic breakdowns -- some would like state-by-state data; failing that, groupings of similar or nearby states would be helpful.
General information about the status of education.	Use of NAEP results as general background information in curriculum and program decisions and in support of policies.	More information on programs, practices and policy information related to achievement.
Information to aid in supporting existing programs and rationalizing new ones.	Use items and objectives in comparisons and in developing state assessments.	More attention to state objectives -- perhaps with some sort of core areas.
Information for use in making recommendations on curriculum and policies.	Use items in areas not commonly measured by standardized tests.	Information on minimal competency, graduation certification, teacher certification, etc.
Effect of school factors on performance.	Use data as general background information and as supporting evidence for changes, new programs, etc.	More items made available to the states -- some would like more innovative ones, some would like more multiple-choice ones.
Information on causes for gains or declines.	State assessment conference.	Stress basics in conjunction with other areas.
Information on current issues related to education.	Use goals and objectives to determine national consensus priorities -- as input to own statements.	National data on compensatory programs; Title I.
Objectives and items for use in state assessment programs, curriculum development.	Use data in considering curriculum changes, modification or as support for existing programs.	Better and more consistent documentation of items in an area so people can see what's available.
Information on Assessment methodology, innovative measurement and scoring techniques.		More small-scale, current-interest areas.
Information on Assessment utilization and dissemination of findings.		Grade data.
Information on goals and objectives in various learning areas.		Definite assessment schedule that states could be informed of in advance.
Exchange of information on state assessment programs.		Better definitions of what is included in change measures.
		Describe other sampling alternatives for states.
		Cooperative developmental efforts between states and NAEP.
		Relationship of instructional variables to performance.
		Continued broad coverage of subject areas.

Use of National Assessment for replication purposes has been dependent on particular state policies and programs. Some states have done extensive replication studies and have used results in making program changes. For example, Minnesota discovered that its students were slightly below the nation in science performance and instituted a task force to study the state's science education program. Connecticut, which replicated Assessment items in citizenship/social studies and career development, established advisory committees to interpret findings and make recommendations about improvement strategies. Texas used NAEP writing and citizenship items in assessing 9-, 13- and 17-year-olds. NAEP procedures were followed so that findings could be used to compare achievement of the three age groups with their contemporaries across the nation. Reports on assessment results were presented to the Texas State Board of Education.

Another important use for Assessment data is to support state findings and to provide objective evidence about the condition of education nationwide. As one official put it, "NAEP data are relied upon heavily to lend support to state findings." Officials from several other states agreed, saying that NAEP data are used to support or corroborate state data and to provide background information. Results are sometimes cited as a rationale for policy changes. Several states said that NAEP data have been cited in legislative hearings or in response to requests for information from the state boards of education.

National Assessment data are also used to put state trends into perspective. As one state official said, "National Assessment is an essential adjunct to the state data because it gives some perspective to the trend information available in the state. If the state and the national data are both going down, this tends to indicate that the causes are probably widespread and related to cultural things rather than specifically related to education practices in the state."

States also use Assessment items and objectives. The increase of objective- and criterion-referenced testing by state agencies has created a great demand for items. NAEP has large pools of items in a number of subject areas -- items that have been used in the Assessment and thus have been, in effect, field tested on a national sample. Data are available for the items so that rough performance comparisons can be made even if Assessment procedures are not replicated exactly.

Several state officials cited the value of National Assessment's measurement of areas other than the basics and its development of innovative measurement methods. "There are virtues in NAEP's demonstrating that assessments can be conducted in other areas," one official said. "We may not be able to afford to do art/music/-literature but at least we have an awareness of what the national achievement is and know that the area can be assessed if need be." Many state respondents felt that NAEP's participation in developing methods of assessing and scoring written essays is one of the more valuable contributions made by the project. One official noted that the use of student writing samples and the use of nonmultiple-choice procedures, specifically the use of interview procedures, were excellent innovations.

Assessment officials were enthusiastic about the Large-Scale Asssessment Conference conducted annually by NAEP in June. One said, "It's an excellent annual conference and cannot be replaced by other conferences." Another called the conference "extremely useful," and yet a third said, "It's a good way to keep in touch with others and find out what's going on in other states. There's no other conference like it around that I know of." It should be noted that the Large-Scale Assessment Conference covers a far broader range of assessment issues than simply National Assessment, focusing on many of the concerns that state and local assessment personnel face in their programs.

Suggestions for Enhancing National Assessment in Accord With State-Level Needs

Most felt that the level of NAEP reporting was adequate for their needs and that communication, especially through the NAEP Newsletter, was good. There were some suggestions for improving reporting of results. An official who uses NAEP information to respond to people who call in with questions felt that more summary information on subject areas, perhaps in the form of brief summaries inserted in the Newsletter, would be useful. He commented that NAEP subject-area reports are often too long and detailed for people who have to deal with large volumes of information.

Several state officials felt that more interpretation of NAEP data would be very useful. One remarked, "Achievement data needs to be put in a context ... some-

one needs to begin thinking about what it might mean. To augment interpretation you need to consider the results of other similar kinds of findings." In a similar vein, another said, "One problem in using NAEP-like results is the difficulty in describing item-by-item data ... NAEP needs to do some interpretation. However, you need to make clear when reporting stops and interpretation begins."

There was also some feeling that NAEP should do more in producing reports that summarize findings across assessment areas. One official suggested further analysis of background variables across assessments to obtain more of a longitudinal flavor of performance in various subject areas. Others supported this view.

Some noted that NAEP reports do not appear until some time after an assessment -- sometimes up to two years after. Although data are accessible, by the time they become available they may no longer be optimally useful.

A request from one state educator is that NAEP provide more information on the content of the exercise pools used to measure change, since without this information it is difficult to determine whether the items used in a particular category or domain are adequate to represent performance within that domain.

National Assessment items serve as a valuable resource to many state assessment programs. Several officials suggested that perhaps National Assessment's release policy should be reevaluated. One commented that although National Assessment says released and unreleased items are comparable, "We only see the released items, and means on these may differ from the means on the total pool reported by NAEP." Another said, "There is suspicion that we don't get the best items." This official would like to see NAEP release all items so he would have a bigger pool to select from. He suggested that NAEP develop a computerized item bank containing NAEP items and items obtained from the states. Another official also questioned the need for the level of security that NAEP maintains on its change items, saying, "Given the fact that in the NAEP samples no youngsters know specifically which items they would be receiving, the question of test security may not be as much of a concern as NAEP tends to think." He added that he would like to see additional items made available from National Assessment.

Assessment personnel complimented NAEP's innovative practices, particularly in the area of writing, and said that greater efforts might be made to publicize those methods. One noted that people are aware of ETS's scoring methods but not of NAEP's and that NAEP might be more aggressive in "marketing" its methodology.

Cost was noted as a drawback to replicating NAEP items and to making comparisons with national results. Commented one official, "Following NAEP replication procedures does mean more cost in printing booklets for each student. Maybe NAEP could do some testing to see if comparable results can be obtained by using separate answer sheets. This would certainly be more usable at the local level." Release of intact item booklets (packages), such as those provided in the Basic Life Skills (BLS) Kit, rather than the loose-leaf released exercise sets presently provided by NAEP would reduce reproduction and assembly costs. Greater use of multiple-choice rather than open-ended exercises would also cut costs; however, others noted that a major strength of the Assessment is its innovative measurement practices, which set it apart from the typical norm-referenced, standardized test.

Some state respondents did feel that more emphasis should be placed on basic skills. The representatives of the National Coonference of State Legislatures noted that, for their purposes, the basics should be stressed although they should be linked to other areas. However, most respondents felt that assessments in separate subject areas are much more useful and should be continued. As one said, "A case could be made for NAEP to concentrate on other than the basic areas because, in my state at least, emphasis is being placed on the basics with state-level programs." Most also supported the continued inclusion of special probes, or small-scale assessments in areas of topical interest. One said, "Access to special probes is good. NAEP saves us a lot of money in this area."

Many people are interested in comparative information on state and national performance. Many state officials said that, while they recognized the limitations of SAT and ACT data, these sources did provide them with state/national comparisons. Such comparisons were often reported in the media and seemed to serve a function desired by the public. Opinion was divided as to whether NAEP should collect and report state-level data. Some felt that it would be very useful given the demands for information in their particular states; others did

not think that the usefulness would justify the probable cost. Some expressed the opinion that collection of state-level data would usurp state functions and responsibilities.

Background variables were a major area of suggested enhancements. State officials were nearly unanimous in seeing a need for finer regional breakdowns, noting that the states presently lumped together in a single region are often unalike. Another official agreed that the regional sample is too diverse for comparison purposes. He suggested perhaps grouping heavily industrialized states or doing additional work in identifying cosmopolitan or larger metropolitan areas to provide a better basis of comparison for his state.

Most of the state assessment people who had attempted to replicate NAEP items felt that grade-level data are much more useful than age-level data. NAEP has available data for modal grades (grades containing the majority of students at an age level -- for NAEP, 9-year-olds in grade 4, 13-year-olds in grade 8 and 17-year-olds in grade 11). Many officials seemed to feel that while modal grade data are not optimal, they are better than no grade data at all.

States that had used NAEP for replication said they tended to use only national, regional and perhaps sex categories. As one said, "We found that we could not readily compare to the National Assessment community-size distinctions because they didn't match our population very well." He added that NAEP might consider a different method of indexing socioeconomic status, perhaps by asking questions that relate to parental occupational levels. He suggested retaining racial identifications and adding questions regarding respondents' linguistic backgrounds.

One assessment director commented, "We're trying to get away from asking students about their background factors and trying to ask more about their attitudes and school experiences. We're also trying to get more information from teachers to describe classroom settings and experiences." Another official also stressed the need for National Assessment to include more information on students' school experiences, such as students' exposure to various curricula and amount of instruction they receive. He commented, "I think this type of context information is very valuable. While it doesn't provide cause-and-effect relationships, it does get people to

think about what might be some possible relationships."

Representatives from the National Conference of State Legislatures made the comment that two things National Assessment claims not to be able to do -- provide specific program information and provide information about the causes of achievement declines -- are the things that legislators are most interested in. They observed that National Assessment is not in the forefront of current issues about education -- issues such as competency testing, graduation certification and teacher certification. In addition, National Assessment information is not sufficiently tied to the development of curriculum materials in their opinion.

One assessment director suggested that more interaction between the states and National Assessment would be beneficial. He suggested that states and NAEP might collaborate in developing some items and objectives so that some of the items and objectives would reflect state goals.

National Assessment has not been particularly active in addressing questions of minimal competency, which presently is a major concern of the states. An official suggested that NAEP might develop a subset of items at a minimum skill level -- items that 90% of the students should be able to answer. National Assessment might also develop a "core indicator" assessment and provide grade-level data so that more states could make comparisons with national-level data.

Although they had many suggestions for enhancing the National Assessment program, the people we contacted at the state level were, in the main, supportive of the Assessment effort. One said, "NAEP is about the only source you can refer people to for items and national data." State assessment personnel were pleased with the accessibility of NAEP staff and felt very appreciative that they could just pick up the phone and generally get information or assistance with problems quickly. Some felt that any type of computerized, direct access system would not be nearly as responsive.

One state official summarized his feelings about National Assessment's role in this way: "NAEP's job is to raise the national consciousness about our educational health. Our educational health includes more than basics. It includes what kids know and can do, as National Assessment has stated it, in the major learning areas."

C. NEEDS FOR ASSESSMENT DATA AT THE LOCAL LEVEL

For any real changes in education to occur, the changes must be implemented at the local level, the place where interaction between student and teacher actually occurs. If innovations, curriculum changes or policy decisions are not communicated to the teachers and supported by them, it is most unlikely that changes in practice will be seen by students. Many innovations imposed from on high, supported by research and eminent education theorists, have failed because teachers either were not taught how to implement a program or had no personal stake in implementing it. Witness the fate of the "new math"!

Serving local needs is of importance to National Assessment for several reasons. First, National Assessment realizes that a large-scale education project, designed to provide information about the schools, must provide information to those actually involved in the process of teaching as well as to those who set education policy and conduct education research. Second, NAEP's data collection takes place at the local level. National Assessment is a voluntary program, and schools and districts have the option not to participate. The amount of testing done in the schools has increased considerably in the last decade, partly in response to the minimal competency and accountability movements. Schools and local districts are understandably reluctant to sacrifice instructional time to the Assessment unless its benefits to them are made clear.

It is difficult to identify a universal set of local needs for information in light of the large number of school districts in the United States, each with its own particular problems and needs for information. The questions needing answers in a small, rural school district may be very different from those faced by large, urban districts. Table 7 lists needs for, utilization of and suggestions for enhancement of the Assessment made by persons at the local level. In considering the needs of local users, ECS drew heavily upon the experiences of NAEP staff members who had worked with local districts and subcontractor staff who conducted the assessment in local districts in addition to talking directly with local school people.

TABLE 7. Needs for, Utilization of and Suggestions for Enhancements of National Assessment: Local Level

Needs	Utilization	Enhancements
General information on national status of education.	Replication of NAEP items for comparison purposes.	Need for grade-level data.
Comparison of local performance with other district, state and national trends.	Use of NAEP items and objectives in curriculum development.	Continue probe assessments (small-scale, current-interest topics).
Effect of school factors on performance.	Uses of NAEP items for program and student evaluation.	Need for more information on programs and their relation to achievement.
Information on policy questions.	Use of NAEP methodology.	More schooling variables.
Objectives and items that can be used in curriculum development and evaluation.	Use as input to program changes.	More regional conferences.
Information on Assessment methodology, dissemination and utilization.	Use to identify problem areas that might need monitoring or attention.	Need for more indexing and categorization of NAEP items, better documentation.
Information on innovative measurement and scoring techniques.		Need for better dissemination of NAEP to locals.
Objectives and items that can be used in instructional programs.		Need to relate NAEP to other studies and to state studies.
Information on how to translate research data into policy options; guidelines on making curriculum changes.		More interpretation of NAEP data in reporting.
Information on relationship of background factors to achievement.		Cheaper methods -- release intact booklets, use separate answer sheets.
		Opportunity to obtain local results.
		Smaller geographic breakdowns -- for example, by state or by county.

Most local school districts gather extensive achievement information about their students, ranging from classroom tests to district and state assessment programs to norm-referenced standardized tests. Much of this information is used to make decisions about individual students. However, districts are also interested in obtaining information for use in program evaluation and improvement.

Although NAEP might be useful to local districts, many are only marginally aware of NAEP'S existence. The 1976 General Accounting Office study surveyed a sample of some 710 school districts, generally questionning either the director of testing or the superintendent of schools, and found that slightly less than half the respondents expressed some degree of familiarity with NAEP. Eleven percent had read some project reports and critiques and 3% had read a substantial number of them. Only 2% of the respondents had reviewed project materials sufficiently to determine its usefulness to their own assessment programs. Of local districts that had districtwide testing or evaluation programs, 90% used standardized testing material.

Utilization of National Assessment at the Local Level

The project has limited the amount of direct technical assistance to local education agencies because of the costs involved. A number of local school districts that have received direct technical assistance from National Assessment either showed potential for demonstrating how NAEP project procedures could be used at the local level or were part of state assessment programs that utilized NAEP. Appendix A lists some of the local districts that have used NAEP rather extensively and shows ways in which NAEP was used.

Several districts have replicated NAEP assessment items and compared their findings to national results. The Cheyenne Mountain (Colorado) Public Schools used NAEP items and procedures to assess students in mathematics and writing. Local results were compared to NAEP national, regional and type-of-community data. Although Cheyenne Mountain did well in comparison with groups reported by National Assessment, performance on some tasks was not as high as expected. As one example, students' persuasive writing abilities were found lacking and consultants were brought in to assist in developing a program in this area. The Lincoln, Nebraska, public

schools also used NAEP items and procedures in assessing writing. Based on results, instructional policy decisions designed to improve writing performance were made and implemented. A mathematics assessment, using input from educators, business leaders and parents was developed by the Roanoke Rapids, North Carolina, schools. About 30% of the items were drawn from the NAEP mathematics assessment, and comparisons to national results will be made.

Others have used NAEP items to assist in building their own assessments but not necessarily to compare their data to national results. Interest is generally higher in areas not traditionally measured well by standardized tests -- citizenship/social studies, writing, career education, art and music. As with states, a major interest in recent years has been in the assessment of writing. As one NAEP staff member said, "If I had the time and money were available, I could spend all my time going to local districts and helping them with writing assessments."

National Assessment objectives and materials are also used in curriculum planning and program building. For instance, in Covington, Georgia, NAEP social studies materials have been used to justify program changes. The district developed a new K-12 social studies curriculum, instituted some new courses and upgraded others. The Boulder Valley (Colorado) Public Schools have also used NAEP citizenship/social studies objectives in designing their own objectives and programs.

It is difficult to determine accurately the extent of NAEP usage by local districts. NAEP mails out a large volume of materials, many of which are sent to local districts. However, National Assessment does not know the extent to which materials are used or how they are used.

NAEP also has value to local schools in less direct ways. Some districts use NAEP as a sort of monitoring system, checking to see whether similar trends appear in their schools. Checks may be as informal as simple discussions by school staff or more formal procedures may be used.

Suggestions for Enhancing National Assessment in Accord With Local-Level Needs

If National Assessment is to be utilized widely at the local level, more local educators must become aware of its existence. As one local educator said, "Not many teachers and curriculum people that I have met have ever even heard of NAEP. You need more popular articles in NEA journals, education journals. NAEP is 10 years old and many school administrators just don't know about it." An educator who conducts workshops for mathematics teachers commented, "Maybe half the people have heard about the National Assessment mathematics results, which have been well-publicized. And the people attending these workshops are the professionally-oriented ones. There is a large group that needs to be reached."

Suggestions for increasing familiarity with NAEP included articles in education journals, increased staff attendance at conventions, presenting exhibits, distributing materials and conducting presessions at conventions, and working with teacher organizations at the state level rather than only at the national level.

It was noted that difficulty in translating research information into action at the local level is not unique to National Assessment. One researcher stated, "There is a big gulf between local school systems and research efforts or the education leadership. The delivery system is not adequate and nothing ever hits at the school levels. We all need to start working with the districts."

A suggestion made was that NAEP might go directly to local districts, perhaps through regional workshops, and show them how results could be used. As one local official said, "Regional conferences are a good forum -- like the NIE/ECS Minimum Competency Conferences a few years ago -- they were excellent." NAEP could conduct workshops on, for example, what the data say to curriculum people and offer to conduct one when districts are considering curriculum revisions. This could become a self-supporting endeavor. There was some concern expressed on the part of a state official that such activities might usurp state functions.

Many suggestions for enhancements at the local level were similar to concerns expressed at the state level. There was general agreement that NAEP needs to do much more interpretation of its data to make them useful to local users. One comment was, "Most school

districts are small and have neither the manpower nor the expertise to interpret NAEP data or use it. National Assessment should be more aggressive in interpreting trends and in offering suggestions on how to combat trends and on what changes to make." NAEP results also need to be related to data from other studies and from the states. A researcher commented, "You need to get to people and lead them through the data. They just can't use it off the cuff." Users would like information on what the data mean and suggestions about what to do about problems that are discovered.

A number of districts make use of NAEP items, both in developing their own tests and in using items to compare their results to the nation. Like state personnel, local personnel want to see more items made available for their use. One official suggested that the indexing and categorizing of items could be improved, saying, "Sometimes items are indexed by theme, sometimes by objectives. There seems to be no standardization."

Officials also wondered if more cost-effective procedures for conducting an assessment could be instituted so that it would be less expensive for them to replicate Assessment items. Some suggestions that were made were using separate answer sheets and releasing entire Assessment item booklets so officials did not have to spend time sorting, selecting and collating items. Of course, complete booklets would reduce a district's ability to construct a unique test, framed in accordance with district goals and objectives. School personnel also felt that NAEP should continue to pay close attention to the balance between open-ended and multiple-choice exercises because while open-ended exercises in some cases provide very valuable information, they are expensive to score.

One official emphasized that more program information would be valuable, saying, "We need more background information to help answer the 'whys' and to provide more information on type of programs in relation to Assessment findings." Grade-level data was also seen as useful, although the modal grade data provided by National Assessment seemed satisfactory to some.

Finer geographic breakdowns would make data more meaningful to local users. It is difficult for a district to make meaningful comparisons to the nation since the demographic composition of any district is likely to differ from that of the nation, and these differences in demographic factors could be used to explain any

achievement differences that exist. Of the local districts included in the 1976 General Accounting Office survey, 65% responded that smaller geographic breakdowns, such as state or county, would make National Assessment more useful to them.

Local districts are also interested in program evaluation data. However, it would be close to impossible and perhaps inappropriate for a national project to be used in evaluating local programs. Program information to be evaluated would have to be fairly general. Gathering data on relationships between instructional variables might, however, be feasible. Local districts might have to conduct some additional studies to see if the information held true in their particular case, but at least they would have been provided with a reasonable starting place for their investigation.

Continuation of probes, or small-scale assessments on topics of current interest, was encouraged. Local districts have found the results useful, since they do not have much information in areas such as health and energy. They also have shown interest in the items, since items in less traditional areas are not always readily available.

Some local districts might be interested in using National Assessment items in building minimal competency instruments. National Assessment has developed materials containing selected items that could be used in assessing minimal competency or basic skills and does make them available to local districts. However, other sources would probably provide more comprehensive sets of items specifically aimed at minimal competency or basic skills.

Local use of National Assessment during the first 11 years of the project's history has been sporadic at best. Some districts have made extensive use of NAEP and have found the results very useful. Often these districts are those with resources for evaluation and with the staff and desire to conduct program and curriculum evaluations. Large urban districts, with a host of problems in addition to student achievement, have been less likely to use NAEP data. Increasing local use of NAEP would require a substantial dissemination and interpretation effort.

SECTION II

WAYS TO ENHANCE THE UTILITY OF THE NATIONAL ASSESSMENT

Consideration of ways to enhance the utility of the National Assessment raises the questions "Utility for what?" and "Utility for whom?" As seen in the preceding section, the needs and suggested enhancements expressed by different Assessment audiences are far from identical. For example, curriculum specialists would like a broad base of data in a subject area, while researchers are far more interested in background and program variables; state and local personnel would like to see more items developed and made available to suit their needs, but accurate measurement of change requires that as many items as possible should be kept secure. More innovative analyses might be useful for research purposes; more interpretation and recommendations related to teaching methodology might be more appropriate for the practitioner.

The potential effect of changes on the Assessment's present primary function -- measuring changes in education achievement over time -- must be carefully evaluated. Extensive resources have been invested in the National Assessment program; much valuable experience has been gained and a substantial amount of data has been collected. Although the program might shift some of its emphases to better meet its audiences' needs, changes should be made in a deliberate manner so the information gathered in the last 11 years remains useful.

As currently structured, the National Assessment combines features of descriptive and analytical studies. Optimizing the Assessment for descriptive purposes can be detrimental to analytic features and vice versa. A descriptive survey demands breadth -- a broad sample, broad coverage of subject areas and objectives within subject areas, many background variables and many different kinds of measures. Since it cannot be determined in advance which features will be of interest over a long period of time, which features will be of interest in the future or which areas will produce changes, it is

necessary to measure as many as possible to provide for adequate description. Analytical studies, on the other hand, often require depth of measurement. Discovering relationships between socioeducation factors and achievement is facilitated by a sample deep enough for fine-cell analysis (analysis of groups defined by more than one variable, e.g., sex X race or region X sex) and by precise background variables. Some of the features that make NAEP a useful descriptive study (breadth of coverage, matrix sampling) are not helpful to researchers interested in longitudinal work, easily collapsible data or experimental control.

A tension also exists between long- and short-term measurement concerns. Policy relevant studies often demand the ability to obtain data quickly on topics of current concern. Many different policy areas might be examined, and those that are of pressing interest today may not be so tomorrow. However, the Assessment has been designed to provide information on long-term trends and has a limited capacity for addressing issues upon which immediate information is needed.

An assessment designed primarily for providing research data would differ considerably from that presently in existence. The National Assessment is not designed as an experimental study but rather is planned to provide survey data about many factors. Relationships among these factors can be explored and used in generating hypotheses about cause-and-effect relationships -- however, the controlled intervention necessary for experimental research cannot be conducted satisfactorily at the national level with a large sample. Cohen (1970, p. 120) suggests that census-like data serve a different function than experimental data.

> By definition a census measures stasis, it quantifies how things stand. If done well, it can reveal a good deal about the interconnection of social structure; if it recurs, it can throw much light on how things change. But no census can reveal much about change other than its patterns -- probing its causes and dynamics requires rather a different research orientation.

Presently, the National Assessment serves a unique function in providing a picture of changes over time in various subject areas for a number of consistently defined groups. If the design of the Assessment were changed to encourage experimental studies, the Assess-

ment might simply become another collection of research studies and the unique information that it provides might be lost. In addition, data collection, as currently structured, is dependent on the cooperation of local school districts. Thus, the needs and desires of local school districts must be taken into account in designing the Assessment. An assessment intended solely for research purposes with many background measures and a sample intended to provide in-depth analysis of group performance and without the ability to provide broad coverage of items and comprehensive measures of change might not inspire local schools' cooperation.

The following discussion examines some of the trade offs involved in making some of the enhancements suggested to the National Assessment program. Some potential enhancements and trade offs are summarized in Table 8. It should be noted that Table 8 is intended to stimulate discussion of potential options for National Assessment. The discussion of trade offs is by no means complete; costs are expressed only in relative terms. Activities designated as feasible within the present model are generally activities that are technically feasible. Political, economic and philosophical considerations have not been fully evaluated.

TABLE 8. Changes in the National Assessment Model and Potential Trade Offs

Suggested Changes	Compatible With Present Model?	Cost	Trade Offs	Problems
DISSEMINATION				
Special dissemination strategies --Targeted mailings to researchers describing data base --Summaries of areas for quick reference --Brief abstracts of reports --Increased reporting in education journals --Increased presentations to professional organizations	Yes	Cost for each strategy is not large; costs for all would not be insignificant	If all strategies cannot be afforded, need to decide which ones will most enhance Assessment utility	Assessment Policy Committee will have to decide which groups are most important to reach and determine optional ways of reaching them
Produce cross-area reports	Yes	Minimal	Staff time for analysis and writing must be made available; may have to draw from other activities	Need to determine how resources should best be allocated
Conduct workshops to assist states and locals in using Assessment data	Yes	Moderate to high	In addition to staff and travel costs, some additional technical support would be needed for analytical and publication efforts	States and locals might expect more information about how problems can be solved or more cause-and-effect data than are currently available
Quicker turnaround in reporting results	Yes to some extent	Moderate to high	Might reduce amount of analysis able to be done	Greater resource allocation to analysis and data processing would be needed
INTERPRETATION AND DATA ANALYSIS				
Provide more data interpretation	Possibly	Depends on extent of interpretation	Resource allocation	What level of interpretation is desired and who should do interpretation? Data may not support the desired levels of interpretation
Design assessments that lend themselves to interpretation	Possibly	Moderate	Some interpretive efforts require modification of Assessment model; may sacrifice some breadth of coverage to answer interpretive questions	Need to determine interpretive directions while assessment is being developed

TABLE 8 -- *Continued*

Suggested Changes	Compatible With Present Model?	Cost	Trade Offs	Problems
Provide advice on intervention strategies with regard to findings	Possibly	Depends on extent of effort	Availability of resources for outside interpretive efforts	Who would do interpretations and what data sources would or should they use in developing advice? Advice might be dependent on interpretive point of view rather than supported by assessment data; as a federally funded agency, should NAEP provide program recommendations?
Conduct more data analyses	Yes	Depends on extent of effort	Reallocation of resources or additional resources would be required: (1) to permit greater involvement of data analysts in assessment development and (2) to cover additional analytical activities	Probably would need to expand data processing and analytical capacity. Might have to draw resources from data collection. Would need to decide on relative quality and efficiency of in-house analysis v. analyses done by outside researchers
ADDITIONAL BACKGROUND VARIABLES				
Collect additional background information --Better socioeconomic-status indicators --Information on language spoken in the home	Possibly	Could range from small to large depending on type and quantity of variables	With a finite amount of time for an assessment booklet, addition of background variables means less time for achievement measures under the present design	Background variables to be included need to be considered very carefully to ensure that they will generate useful information
Collect data on income/ parental occupation to make compatible with SES data in other studies	Possibly	Depends on procedures used -- could range from small to very substantial	Difficulty of collecting reliable information at reasonable cost	Information probably would have to be collected from students and it is difficult to know how reliable this information might be; might not be allowed to ask students about sensitive topics

TABLE 8 -- *Continued*

Needs	Compatible With Present Model?	Cost	Trade Offs	Problems
Collect background data compatible with other studies	Possibly	Depends on extent	Might have to sacrifice some variables that have been used in measuring change or reduce achievement measurement time	Criteria for deciding which studies to become compatible with. May have difficulty in collecting information from students and schools
Collect data on program and instructional variables	Possibly	Additional costs involved -- depends on methods used	Larger sample or special supplementary samples may be needed to get at program variables; resources	Data on programs can be collected at school level but definitions may vary nationwide. Many programs are not national in scope or affect too small a group of students to be picked up in the present sample. Under present design, cannot ascribe causation to relationship
Collect more school-level data	Possibly	Additional costs	Would require extra time on the part of school officials; many of the variables desired might require observations or interviews by field staff	Present sample of schools probably too small to provide meaningful information. Would be difficult to tie school-level information to achievement under present design
Use finer geographic breakdowns -- census region, state, county, district	Some census region data are available; other breakdowns are not possible under present sample design	Increasing sample size involves substantial added costs	Resources; may need to sacrifice some breadth for depth	Finer geographic breakdowns are desired. How fine breakdowns should be and whether this capacity should exist at summary or exercise level need to be explored. Schools, districts or states may not wish to be identified; may refuse to participate in assessment
RESEARCH ENHANCEMENTS				
Packaging of items to permit construction and analysis of subscales within an item booklet	Possibly	Some additional costs	Packaging strategies necessary for accurate measurement of change might be adversely affected	Research needs would need to be determined at the developmental and packaging stages

TABLE 8 -- *Continued*

Suggested Changes	Compatible With Present Model?	Cost	Trade Offs	Problems
Expand sample to permit study of other groups (e.g., other racial/ethnic groups, bilingual groups)	Yes	Substantial	Would probably have to limit achievement measures to compensate for increased sample cost	Some groups are too small to survey efficiently with the NAEP sample design. Efforts might not be well-received by ethnic groups involved
Change project design to make more useful for program evaluation, experimental research	Very limited	Substantial	Would probably lose capacity to measure change and lose broad-based achievement measures	Loss of utility of previous data base. Little of present model would be efficient or appropriate for this purpose
POLICY RELEVANT ANALYSES				
Use of policy relevant sampling units to match justificational units	See finer geographic breakdowns			
Ability to include policy relevant questions	Possibly	Cost would vary with type of questions to be addressed	Might have to give up some achievement questions or restrict ability to measure change	Need mechanism for selecting questions that are important and that will remain of interest for the length of time required to collect data. Need way to include them in the assessment in a cost-effective way. Same difficulties as with background variables
Add policy relevant areas; add more subject areas	Yes	Substantial	Would need to make sure information was desired by interested groups since other areas might have to be sacrificied. Might be difficult or meaningless to measure change in policy areas; might detract from present measurement of change	How to determine policy relevant issues, how to decide how much emphasis to give them. Need for shorter analysis and reporting times in many cases

TABLE 8 -- *Continued*

Suggested Changes	Compatible With Present Model?	Cost	Trade Offs	Problems
Relate policy relevant questions to achievement, e.g., length of school year, school size, work experiences	Possibly	Moderate	Would probably hamper measurement of change; would need space for policy questions; relating to achievement might require larger sample	Would require increasing response burden on school officials
STATE AND LOCAL NEEDS				
Release as many items as possible	Perhaps	Moderate to large	Uncertainty about effect of lack of item security	Might jeopardize measurement of change; might require development of more items
Develop objectives and items related to state goals	Probably not	Fairly large	Would make data more useful to states; might hamper measurement of national change	How to determine which states' goals to use; what percentage to use. Question of adequacy of "nationwide" perspective
Change sample design to allow state- or finer-level reporting	Yes	Substantial	Resources, cooperation problems	Would allow better comparisons; might violate confidentiality; might be politically unacceptable. Schools, districts or state might refuse to participate in assessment if they felt there was any chance they could be identified. Would have to investigate very carefully and study potential benefits compared to potential drawbacks
Conduct workshops to help state and locals use data, materials and methodology	Yes	Moderate to large	Would require substantial staff time and involvement; special materials would probably be needed; resource reallocation would be required	Present capability to interpret materials might not provide states and locals with the kind of information they want

Dissemination and Data Accessibility

Many people, even within the education community, are not aware of National Assessment. A more concerted effort to publicize the Assessment and ways that it can be used would be useful.

Most audiences saw a need for more and better dissemination of Assessment materials. Several people suggested that, while the newsletter and general reports are suited to the audiences they serve, short summary materials for those who deal with large volumes of information and special communications targeted to researchers are needed. Researchers also need more information on the data base and its analysis potential. At the local level, workshops were suggested as perhaps the most effective way of reaching educators. Other potential ways of disseminating NAEP information are listed in Exhibit 1.

EXHIBIT 1. Ways of Providing Information

- News releases
- News conferences
- National, state, local conferences
- Direct mail announcements
- Data tapes
- Item collections/banks
- Objectives booklets
- Guidebooks and "how to" manuals
- Newsletters
- Television and radio appearances
- Education journal articles
- Printed reports
- Filmstrips
- Microfiche
- Speeches
- Workshops and seminars
- On-site visits
- Telephone
- Correspondence
- Testimony
- Issue and discussion papers
- Technical bulletins and monographs

Firmer scheduling of assessments would be beneficial to many groups. States wishing to conduct comparative studies would be aware of upcoming assessments; researchers could determine whether forthcoming information would be of use to them; federal agencies and Congressional staff would be better able to make use of the data.

Improved accessibility of data and materials was also seen as desirable. However, the needs of different groups for data and materials vary. State and local assessment personnel would like clearer presentation, documentation and classification of NAEP items. Classification schemes should be as uniform as possible across subject areas, and more documentation of items should be included in general reports.

Data accessibility has been generally sufficient to meet state and local needs since the staff of the User Services Department has been highly efficient in disseminating the needed information. Research use often demands extensive data processing support. Since preparation of data tapes and other materials has not been an ongoing process, the initial production costs have been substantial. More cost-effective production will be investigated in the future.

Increased Interpretation and Data Analysis

Many people suggested that the Assessment would be much more useful if it made more of an effort to interpret its data. The question of interpretation has always been a difficult one for National Assessment. Sensitive to fears about federal control and federal standards in education, the Assessment in its early days avoided making any statements about what the data meant since this might have been construed as establishment of federal standards.

Recognizing that users need interpretation of the data, National Assessment has contracted with various organizations -- including the National Science Teachers Association, the National Council for the Social Studies and the National Council of Teachers of English -- to conduct interpretive studies. In addition, the National Council of Teachers of Mathematics prepared monographs on the results of the 1972-73 mathematics assessment, and a group from the American Personnel and Guidance Association completed a comprehensive publication on responses to career development items. These last two

studies examined item-by-item responses and attempted to provide diagnostic information for practitioners.

However, the extent of interpretation done has been limited. National Assessment can identify problem areas but cannot advise people on how to solve the problems. For instance, in mathematics, deficiencies were found in problem-solving, and it was suggested that more attention be paid to this area, but specific strategies for teaching problem-solving were not provided. Increased background variables may make it possible to discover relationships between education practices and achievement, but the causal links supporting recommendations of one practice or another would, under the present design, be tenuous at best.

National Assessment will continue to pursue the questions of who should interpret results and what level of interpretation is appropriate. Through experience, National Assessment has learned that assessments are most useful if they are designed with specific purposes in mind. For example, the 1977-78 mathematics assessment was designed to answer certain questions about mathematics achievement, and those questions usually could be answered in reporting. Other assessments have been rather generally designed to measure subject-area objectives, and other classification schemes have been imposed during analysis and reporting. Interpretation of these assessments often has not been as successful.

Increased and more innovative data analyses by the Assessment were recommended as a way of demonstrating the potential of assessment data. Some possibilities for analysis strategies are discussed in Issue Paper #7 -- Data Analysis. However, in conducting data analysis, the Assessment cannot lose sight of its responsibility to collect, analyze and report baseline and change data in as timely a manner as possible. It may be that some analytical functions could be more appropriately carried out by other researchers. The results of the National Science Foundation studies on using NAEP data for secondary analysis and the joint NIE-ECS contracts for research to be awarded in 1980 should provide more information on NAEP's potential for research use.

It should be noted that increased data analyses might also demand increased interpretive efforts. For example, correlations might be helpful to the researcher but would require interpretation for other audiences.

Increased Background and Program Variables

Most people felt that more and better background-variable information should be collected and, if possible, related to achievement. Many audiences expressed a need for information on variables that could be manipulated in addition to information on factors that are not amenable to change. Problems related to collection and analysis of background variables will be discussed more completely in Issue Paper #7 -- Data Analysis. Some of the issues will be briefly touched upon here.

National Assessment has collected some program information -- for example, on students' course background -- but it is difficult to collect meaningful program information given National Assessment's current design. National Assessment's sample is designed to provide information at the student level. To relate students' achievement to other factors, ideally the students should provide the information on background or program factors, and their information, particularly with regard to programs, is unlikely to be very reliable.

In addition, each student does not respond to every item booklet. To relate student background information and performance for the entire sample of students and items assessed, the same background questions should be repeated in every booklet, reducing the amount of space available for achievement questions.

Program information can also be collected at the school and district levels. However, it is difficult to be sure this information is consistent because, for example, the same program names may mean different things in different areas. It is also difficult to link information gathered at the school or district level to student achievement. Students from a school are sampled and are not selected according to whether or not they have participated in a particular program. Provisions for including students who had participated in particular programs might have to be made.

While NAEP can provide some information on the relationship between the availability of programs and student achievement, it is not a particularly useful tool for program evaluation. Program evaluation can be better done in small-scale studies where factors can be controlled. The Assessment provides a picture of achievement at different points in time but cannot isolate factors that cause changes.

A number of respondents felt that NAEP background information should be made more compatible with those from other studies. Income data, like that collected for the Census, is one option suggested. It was suggested that NAEP devote more resources to determining optimally useful background variables.

Finer geographic breakdowns were recommended by national education organizations, state personnel, local personnel and researchers. National Assessment presently has the capacity to report data on some finer breakdowns -- for example, the nine census regions, the 10 federal regions or groups of four to five states. However, generally for finer breakdowns, data can only be reported as summaries over groups of exercises; data for individual exercises usually are not reliable. Often, however, state and local personnel need data on individual exercises for comparison purposes. State-level data presently collected could be made available to the states; however, the NAEP sample presently is not necessarily representative at the state level and would have to be very carefully explained if it were made available.

National Assessment is dependent upon the cooperation of local schools for its data collection. Many local districts or schools are strongly opposed to any reporting that might identify schools or districts assessed, since unfair comparisons might then be made. In considering providing finer geographic breakdowns, confidentiality agreements and the concerns of state and local districts must be kept in mind.

Increased Research Potential: Sample and Package Design

Although experimental research and program evaluation are not well-suited to the present Assessment design, many secondary analyses are possible. Some enhancements to aid researchers, such as the addition of more or different policy-related or background variables as discussed above, might not interfere greatly with the present design. Other enhancements must be considered carefully in light of the effect they might have on the Assessment's ability to measure change in performance.

NAEP presently uses a broad probability sample of age groups with oversampling of certain subgroups. No one student responds to all items; each item booklet is administered to approximately 2,000 respondents. When

responses to items given in different booklets are summarized, the sample size is increased accordingly; standard errors are reduced and the estimate becomes more stable. Extremely stable estimates are very important in the measurement of change, since changes over a five-year period may be quite small and, if estimates are not very accurate, may not be discernible due to large sampling variability. Researchers would like packages (booklets) that are designed to answer research questions with a number of items on the same topics. Several researchers noted the difficulty of constructing achievement subscales within a package. Thus the spreading of items on the same topic or objective across booklets to obtain stable estimates may work against optimum research design. Some compromise between these needs might be possible. For example, several packages might be set aside for research use.

NAEP's present sample does not permit extensive fine-cell analyses. National Assessment has generated data for some two-way cell analyses (e.g., sex X race, region X race) but does not have a sufficient sample size for three-way analyses. Two-way analyses on small groups, such as Hispanos, often are not possible. In order to obtain a sufficiently large sample, two-way analysis usually must be done for groups of exercises, not individual exercises. Data on special groups, such as Indians or Orientals, would require expansion of the sample with specific oversampling of those groups. In-depth sampling of particular groups might restrict National Assessment's national sample, if additional funds are not made available.

Research use is also hampered by the long lead time required for an assessment. A researcher who participated in the research design for an assessment might have to wait several years before data become available. To date, there has been little opportunity for researchers' input into the development of assessments. The existing items, data and packaging schemes have been manipulated by researchers after the fact. If greater research usage is desired, it might be helpful to develop a mechanism to encourage research input near the beginning of the developmental process.

Increased Policy Relevance

Representatives of a number of audiences felt that the National Assessment should attempt to become more policy relevant. The question of policy relevance is a many-faceted one, since many different kinds of information might be relevant to policies and many different policies might potentially be examined. The Policy Committee would have to consider carefully the claims of various interests in determining which policy relevant issues to address.

ational Assessment has been useful in encouraging people to think about policies in the areas of mathematics and writing. The data have been used as a source of background information that has stimulated further exploration of problem areas discovered. National Assessment has also done probes, or small-scale assessments, in areas of more topical interest than the traditional curricular subject areas -- health, energy and consumer education.

However, National Assessment would have difficulty in responding very quickly to policy questions requiring immediate answers. Approximately two years are necessary for printing of materials and preparation of tapes, data collection, scoring, analysis and reporting. Development of items requires additional time. Since the three ages are assessed at different times in the school year, data collection for an assessment requires approximately nine months.

There are a number of long-term policy questions that the Assessment might address. NAEP could collect information on such questions as school entry ages, length of the school year, the extent of youth labor force participation, school size, or subgroups such as high and low achievers. In addition to collecting and reporting such information, some of this data might also be related to achievement, although some of the difficulties mentioned under Background Variables would be encountered. The Assessment could also collect data on topics of current interest such as parenting, ecology, ethical and moral concerns, or knowledge about modern technology -- with the limitation that issues would have to be identified two to three years before data were expected to be available. At the federal level there is some interest in comparing private and public school performance. However, this might lead to unfair comparisons or misinterpretation of data. The implications of this type of reporting must be carefully considered.

The Assessment is limited in the types of policy relevant questions that it can ask. As a national study, the Assessment cannot ask questions that might be construed as an invasion of privacy; in addition, schools, districts or states have the right to delete questions from the assessment that they view as inappropriate for their students.

The Assessment is simply too large to gather some types of policy-relevant information. For example, most program evaluations would be better conducted with small studies specifically tailored to the program to be evaluated. Methods of increasing National Assessment's policy relevance must be carefully studied so that the Assessment can be used to answer those questions that it is best suited to answer.

State and Local Needs

Enhancements to serve state and local needs often seem antithetical to needs expressed by researchers. State and local education agencies are interested in data on trends in performance, in the ability to compare their achievement to national levels and in obtaining items and objectives for use in or adaption to their own programs.

Data on a variety of subjects and on various objectives within those subjects can be used by states and local districts to reexamine their own programs. Many are not particularly interested in subgroup data, aside from region, sex and grade, because NAEP subgroups often do not match their particular situation and because they may not have large populations of the racial/ethnic groups surveyed by NAEP. Grade data would be useful for most school personnel because their testing and assessment programs are done by grade. Assessing by grade might be easier and less costly for NAEP since then intact classrooms could be assessed. However, data from intact classrooms could well be affected in various ways by the classroom situation, and this might make Assessment data less reliable.

States and local districts make use of NAEP items in their own assessment programs. Several people suggested that NAEP reconsider its release policy to make more items available. Many felt that NAEP's sampling procedures make it nearly impossible for people to teach to the test, so change measures would not be affected.

National Assessment has taken the position that releasing items would provide a source of uncertainty about the accuracy of measures of change.

Some states and locals would like to have data on which to base comparisons -- state-level or district-level data. Many do use ACT and SAT data for comparative purposes even while recognizing their drawbacks. Others felt that the Assessment should not attempt to provide data at these levels, citing the danger of taking over state and local functions and the danger that confidentiability would be violated. Providing data at the school, district or state level might well endanger the voluntary participation on which NAEP's data collection is dependent. Any changes in the direction of state- or district-level data collection would have to be very carefully studied.

An opportunity to make their concerns known and to participate in the development of assessments might be useful to states and locals. Often, National Assessment objectives do not match state objectives, and states might like an opportunity to make Assessment items more relevant to their needs. A current pressing issue for many states is minimal competency. While National Assessment does not assess "minimal competency," since other instruments are designed exclusively for that purpose, some attention to developing items that might be used for minimal competency could be considered.

States and local districts have difficulty in using research findings. Additional help from National Assessment could be beneficial in assisting state and local agencies to use Assessment data. However, providing more assistance to state and local districts would take a substantial commitment of time and money and would necessarily draw resources from other areas.

Many states and locals would like to see Assessment procedures made simpler and cheaper so that they could do replications at less cost. Use of answer sheets, release of intact packages and more multiple-choice items would be helpful. National Assessment is presently experimenting with the use of answer sheets. Release of intact packages, given National Assessment's current packaging and release strategies, probably would not provide particularly meaningful sets of items. In addition, people would not be able to pick the items they want to use and the Assessment would become more like other standardized tests. The Assessment is moving to the use of more multiple-choice items, but, again,

this strategy jeopardizes the Assessment's ability to provide unique information and to measure areas that are not usually measured successfully using multiple-choice items.

SECTION III

SUMMARY

The National Assessment was designed to serve a unique purpose -- to monitor trends in education achievement in subjects typically learned in school. The Assessment was specifically designed to provide accurate estimates of change in performance over broad subject areas; it differs from norm-referenced, standardized tests in that it measures achievements of groups of students on many different tasks rather than ranking students' performance on a fairly homogeneous set of measures.

Attempts to redesign or restructure the project must carefully evaluate the effect that any changes will have on the functions that the Assessment now serves. Most of the people contacted by ECS and by Abt Associates Inc. felt that National Assessment provides a valuable service. Although most wanted it to do more than it is presently doing, the majority did not want to lose the data and services currently available.

The design of a research project is the main determinant of the results that will be produced. Desired results must be taken into account when designing a project such as the National Assessment, since using a design intended for one purpose and adapting it to other purposes is likely to meet with limited success. Designing a project to attempt to meet all of a large variety of needs that might occur may be a self-defeating strategy; it is generally far more efficient to have definite goals and to devise methods specifically intended to meet those goals.

Thus, in considering the design of the National Assessment for the 1980s, Assessment planners must determine the project's primary goals and primary audiences to be served. Even if resources were unlimited, the Assessment could never hope to serve all its potential audiences since different needs and uses demand different assessment designs. And in fact, since resources are not unlimited, the Assessment presently cannot fulfill all the tasks that it might undertake under

its present design. As one educator remarked, "There are so many tasks and you have to set priorities. You don't want less quality in dealing with priorities if doing all the other things decreases the quality of effort in critical areas."

The American education system is currently undergoing extensive examination -- both from without and from within. Decisions affecting the schools affect millions of young people -- young people whose education is critical to the nation's future. Much more information is needed on how schooling "works," so that decisions can be based on more than intuition or theory. The National Assessment has formed part of the data base of information that we have about the schools and their workings in the 1970s -- a data base that provides a foundation onto which new findings can be added and integrated in the coming decade. It is hoped that the Assessment planners will take into account the many diverse needs for information and the many potential uses for data in structuring the National Assessment so that it can be a valuable resource in meeting the informational needs of education in the 1980s.

APPENDIX A

SELECTED EXAMPLES OF USES THAT HAVE BEEN MADE OF NATIONAL ASSESSMENT DATA AND MATERIALS

National

Congressional Testimony

Before the House Subcommittee on Elementary, Secondary and Vocational Education:

Oct. 1979	Oversight Hearing on Mathematics Achievement. Called following release of NAEP mathematics reports to investigate the seriousness of the problem and discuss possible remedies.
May 1977	Hearings on Testing. At hearings on the testing situation, Dr. Forbes provided an overview of NAEP results as they related to the "declining test scores concept."
Winter 1975	Provided testimony on results of the Functional Literacy Assessment as it related to support of the Right-to-Read legislation.
Winter 1974	Provided testimony on applicability of the NAEP program in the implementation of the Quie amendment concerning federal resource allocation for ESEA on the basis of the educationally disadvantaged rather than the income disadvantaged.

Before the Senate Education Subcommittee:

Spring 1977	Testified regarding possible models using National Assessment in a voluntary national-level testing effort to obtain information on minimal competencies.

Provided information to Congress and Congressional staff:

Spring 1977	Provided Congressman Quie with a technical response to his proposed ESEA legislation, which suggested that NAEP should be used to determine the educationally disadvantaged.

Have conducted Congressional briefings before release of many major reports.

Studies for Federal Agencies

Spring 1980	Department of Labor. Contracted with ECS to conduct study of out-of-school 17-year-olds, using NAEP reading/literature, career development and basic life skills items. Some items specific to DOL informational needs are also included.
1979-80	Department of Defense Dependent Schools. DODDS awarded ECS a contract to assess writing skills of 9-, 13- and 17-year-old DODDS students. NAEP items and procedures were used; findings will be compared to NAEP results. DODDS is also conducting an assessment of reading/literature concurrently with NAEP'S assessment in this area; results will be compared to NAEP's.
1977-79	National Institute of Education. ECS received a grant to conduct the Women in Mathematics survey. The survey of 1,300 13-year-olds and 1,600 high school seniors was added on to the already existing annual NAEP assessment.
1977	Food and Drug Administration. The FDA contracted with ECS to administer several FDA-developed questions in the NAEP young adult assessment. Items included questions on the ability of young adults to interpret information on labels for drugs and food.

Sept. 1977	DHEW/Basic Skills Group. Provided reading and mathematics materials. Materials will become part of a test file to provide current information about assessment instruments to educational researchers.
August 1976	Institute for Responsive Education. Provided data collection and research assistance to IRE in completing the project "A Study of Citizen Participation in Education Decision Making." The project was funded by NIE.
1974-75	National Right-to-Read Program. A Mini-Assessment of Functional Literacy was conducted by ECS under agreement with the National Right-to-Read program. The survey used NAEP reading items and procedures and involved approximately 5,200 in-school 17-year-olds. The survey was conducted concurrently with the regular NAEP 17-year-old assessment.

Dissemination Efforts With Federal Agencies

Summer 1979	Office of Consumer Education. Contracted with ECS to report the findings of the consumer assessment. NAEP produced a brochure and released findings in Washington. Four seminars were conducted for representatives of consumer agencies, education and federal agencies, and congressional staff. The findings were widely disseminated through the media and by consumer groups.
1976	National Science Foundation. Four university faculty members spent 10 weeks conducting special analyses of existing NAEP data. They explored relationships between background variables and achievement and examined methods of constructing an index of basic skills.
1974	Office of Career Education. A grant was awarded to develop a brochure to

	summarize the findings of the career development assessment and to develop an assessment kit for career development to serve as a resource material for persons in career education faced with assessment or evaluation tasks. Materials were disseminated by the Office of Career Education and by NAEP.
Ongoing	National Center for Education Statistics. National Assessment provides, on request, material for inclusion in The Condition of Education , published by NCES.

Activities Involving National Education Organizations

National Assessment has conducted numerous briefings and given many presentations to members of national education organizations.

1979-80	National Council of Teachers of Mathematics. Have used NAEP data as a source in constructing mathematics curriculum for the 1980s. Currently working under a grant from NSF to write articles interpreting National Assessment mathematics data for NCTM journals.
1975-80	National Council of Teachers of English. Have worked closely with National Assessment in developing methods for scoring essays.
1976	International Reading Association. Cite NAEP data as source of information about reading achievement. Established an interpretive team to review, critique and interpret the NAEP reading data released in September 1976. Articles on the implications of the data were written for publication in the IRA Reading Aid series.
1974-76	American Personnel and Guidance Association. Conducted an interpretive study to describe the results of NAEP's assessment of

	career and occupational development. Five monographs were produced and disseminated.
1975	John Mellon of the National Council of Teachers of English wrote a monograph on how first assessment results in reading, writing and literature could be used.
1974	National Council for the Social Studies. A committee of 40 studied and interpreted the 1969 citizenship and the 1973 social studies results. To report their interpretive work, NCSS published a study titled National Assessment and Social Studies Education .

Activities Involving Foundations

1977	Ford Foundation. Grant to conduct a study of the effective utilization of education assessment data. The study compared data gathering and dissemination techniques in the fields of medicine, agriculture and economics.
1977	Wingspread Conference on Global Literacy. NAEP staff participated in conference, examining steps NAEP might take in helping to assess global literacy.

Textbook Publishers

1979	Ginn and Company. Is using NAEP Hispanic data in structuring textbooks.

Research Use

1980	Group from University of Illinois at Chicago Circle investigating potential of NAEP for secondary analysis under a grant from NSF.
1980	Joint NIE and ECS contracts to be awarded to researchers to work with NAEP data.

1980	Nancy Burton, University of North Carolina. Using 1975-76 mathematics data to investigate "balancing" of black/white data.
1977	Ruth Jones, University of Missouri at St. Louis. Used NAEP data in investigating political socialization.
1977	Steven Jurs, University of Toledo. Used 1975-76 mathematics data to investigate differential achievement of males and females.
1977	James McPartland, John Hopkins University. Used career development data in studying occupational socialization of youth.
1979	Ina Mullis, National Assessment of ducational Progress, University of Colorado. Doctoral dissertation relating effects of home and school to 1975-76 citizenship/social studies.

Regional

1980	Clearinghouse for Applied Performance Testing, Northwest Regional Educational Laboratory. Produced monograph of current state of art in assessing writing and methods being used in various states; describes influence of NAEP methodology.
1980	Clearinghouse for Applied Performance Testing, Northwest Regional Laboratory. Conducting national seminar on writing and writing assessment in conjunction with National Assessment at the conclusion of the Large-Scale Assessment Conference in June.
1978-79	CEMREL, Inc. A guidebook to help elementary school teachers increase their students' abilities to express ideas and communicate more effectively

	was developed through the joint effort of CEMREL, Inc., a St. Louis-based educational laboratory, the University of Iowa's Advanced Writing Program and National Assessment.
	State
Ongoing	Large-Scale Assessment Conference. In June, ECS conducts an annual conference for persons involved in large-scale assessments. This provides an opportunity for persons involved in assessments to exchange ideas and to become informed on the progress of assessment in other programs and locations.
1975	Council of Chief State School Officers. NAEP staff met with CCSSO's Committee on Citizenship Education to provide input in the area of citizenship. As a result of this meeting, the CCSSO produced a citizenship objectives booklet that could be used by the states as a guideline in the area of citizenship.

Replication

Over the years, most state departments of education have used NAEP materials and assistance in varying degrees to implement their own assessment programs. While some states have adhered to a strict NAEP replication assessment model (objectives, items and procedures) so that subsequent data comparisons could be made, others have adapted NAEP items and procedures to compliment their own assessment models, making data comparisons on more of a "relative standing" basis. A summary of the type of uses made of NAEP materials and procedures is shown below.

A. States that have closely replicated the NAEP assessment model (items and procedures): Connecticut, Florida, Hawaii, Louisiana, Maine, Massachusetts, Minnesota, New Hampshire, Rhode Island, Tennessee, Texas and Washington.

B. States that have used NAEP materials but adapted items/procedures to compliment their own assessment model: Alabama, Arkansas, California, Colorado,

Delaware, Illinois, Indiana, Iowa, New Jersey, North Dakota, Ohio, Oregon, Wisconsin and Wyoming.

C. States that have received assistance from NAEP in terms of consultation, technical advice and/or materials but no direct evidence of having used NAEP items: Kansas, Michigan, Missouri, Nebraska, New York, North Carolina, Pennsylvania, South Carolina, South Dakota, Utah, Vermont and Virginia.

Selected illustrations of states' uses of NAEP materials and procedures from each of the above categories are as follows:

Arkansas

Arkansas has completed two statewide assessments. The most recent assessment (1977-78) surveyed ages 9, 13 and 17 in the areas of math and citizenship/social studies. All items came from NAEP's item bank and Arkansas followed NAEP's procedures quite closely. Local school districts had the option to piggyback on the state survey. Assessments conducted by Arkansas using NAEP items follow: 1976-77 -- reading; 1977-78 -- math, citizenship/social studies.

Connecticut

The Connecticut assessment program during 1979-80 looked at math and science. The state sample of students to be assessed was based upon age/grade, i.e., 9-year-olds in the 4th grade, 13-year-olds in the 8th grade and 17-year-olds in the 11th grade. LEA "piggyback" assessments were optional on a voluntary basis. As in prior assessment years, Connecticut followed a NAEP replication model and is interested in making national and Northeastern region comparisons. Other assessments conducted by Connecticut follow, as well as a schedule: 1971-72 -- reading; 1974-75 -- science; 1976-77 -- math; 1977-78 -- career and occupational development, citizenship/social studies; 1978-79 -- reading; 1979-80 -- math, science; and 1980-81 -- art, music.

Delaware

The 1979-80 assessment of writing (13-year-olds only) will use NAEP items, procedures and scoring methodology. Writing items will be scored for mechanics (spelling, punctuation, grammar, etc.) as well as pri-

mary traits. Delaware's use of NAEP items is: 1972-73 -- science; 1976-77 -- math; and 1979-80 -- writing.

Hawaii

A sample of students in grades 4, 8 and 11 were assessed in writing in early May 1979. The items were selected by the Northwest Regional Educational Laboratory and a committee of the University of Hawaii and Department of Education specialists. NAEP staff provided guidance, particularly on the scoring process (primary-, secondary- and tertiary-trait scoring). Prior assessments using NAEP items and procedures are: 1974-75 -- reading, math; 1976-77 -- reading, math; and 1978-79 -- writing.

Illinois

Illinois assessed three areas in the spring of 1978 -- math, reading and citizenship. A sample of students in grades 4, 8 and 11 were tested. Some NAEP items were used intact, while others were changed or modified. Illinois plans to assess math and reading during the next three years with a third area being added each year. For the purposes of comparing to NAEP, the 9-, 13- and 17-year-olds are pulled from the grade sample. Following are assessments that used NAEP materials: 1975-76 -- reading, math; 1976-77 -- science; 1977-78 -- reading, math, citizenship/social studies; and 1978-79 -- reading, math, citizenship/social studies, science.

Louisiana

The Louisiana Department of Education conducted a statewide writing assessment for the first time during the 1979-80 school year. The assessment consisted of objective-type items in writing and actual writing samples. NAEP's primary- and secondary-trait scoring methods were used. Prior assessments using NAEP items and procedures are shown as follows: 1973-74 -- reading; 1974-75 -- math, citizenship/social studies; 1975-76 -- career and occupational development; 1976-77 -- science; 1977-78 -- reading; and 1979-80 -- writing.

Maine

Maine was operating under a 1977 state mandate to test all 11th graders and a sample of 8th graders in basic skills (reading, writing, math). Test administration took place in the fall of 1977, with reporting to the state legislature in early 1978. NAEP items were

used almost exclusively. Results were reported for each item, showing comparisons to previous assessments and NAEP national and Northeastern percentages. Other assessments conducted by Maine are: 1971-72 -- citizenship, writing; 1972-73 -- reading, science; 1973-74 -- reading, writing; 1975-76 -- math, career and occupational development; 1977-78 -- reading, writing, math; and 1979-80 -- nutrition -- adopting NAEP procedures.

Massachusetts

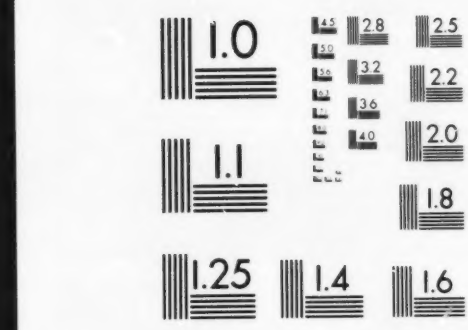
The 1975-76 writing assessment was given to a representative sample of 9-year-olds and 17-year-olds. Massachusetts also conducted a 1975-76 reading assessment (9- and 17-year-olds) and a career and occupational development assessment (13- and 17-year-olds) following NAEP's assessment procedures and schedule. Massachusetts is interested in national and Northeastern region comparisons. Local districts have the option to conduct "piggyback" assessments. Other assessments that used NAEP materials and procedures follow: 1971-72 -- citizenship, science; 1973-74 -- reading, math; 1975-76 -- writing, citizenship/social studies; 1976-77 -- sciene; 1977-78 -- reading, career and occupational development; and 1979-80 -- speaking and listening (being developed through a contract with NAEP staff).

Michigan

In the fall of 1978, the Michigan Educational Assessment Program administered locally developed writing tests on a statewide pilot basis. Because these items were not scorable by the NAEP primary-trait method, Michigan is considering the use of actual NAEP 1978-79 writing assessment items to once again conduct a statewide pilot during the 1979-80 school year.

Minnesota

Minnesota has been a heavy user of NAEP items. Approximately one-third of each their assessment's items come from NAEP's released items. For a couple of years, Minnesota did not conduct any full-scale, statewide assessment programs. Rather, all efforts were directed at helping local districts set up their own testing programs, drawing upon instruments that had been developed previously by the state. During 1977-78 the state department once again began developing instruments and assessing to get state-norms so these materials could be used by local districts. While the state does not plan to assess basic life skills during 1979-80 for state



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

results, they are making the NAEP items available to local districts. Plans are underway to use NAEP's art, music and literature items for a 1980-81 assessment. Following are assessments that used NAEP materials: 1972-73 -- reading; 1974-75 -- reading; 1977-78 -- citizenship/social studies; 1978-79 -- math, science; 1979-80 -- writing, basic life skills; and 1980-81 -- art, music, literature.

New Hampshire

The NAEP Bicentennial Citizenship Kit was used by New Hampshire to assess a state sample of about 500 13-year-olds and 500 17-year-olds.

North Dakota

North Dakota has conducted math and reading assessments using NAEP items where appropriate and following NAEP's schedule for assessment -- 17-year-olds during the spring, 13-year-olds during the fall and 9-year-olds during the winter. Their three assessments follow: 1976-77 -- reading; 1977-78 -- math; and 1978-79 -- math.

Ohio

An English composition assessment was administered to 8th graders in 1977 and to 12th graders in 1978. Both tests consisted of a combination of essay items and multiple-choice items. The objectives paralleled NAEP objectives, and several NAEP items were used to allow for primary-trait and mechanics scoring (some items were also scored for secondary traits). A schedule for other assessments using NAEP items follows: 1973-74 -- reading; 1975-76 -- reading; 1976-77 -- reading, writing; 1977-78 -- math; and 1978-79 -- math, citizenship/social studies, science, basic life skills.

Oregon

NAEP procedures and materials were the focus of a statewide writing assessment during the spring of 1978. A statewide sample of about 3,500 students in each of three grades (4, 7 and 11) participated. In addition to the NAEP items, a few locally developed items were used, but NAEP's scoring procedures (holistic) were followed for the analysis. The writing exercises were scored and results returned to students.

Rhode Island

Rhode Island uses a norm-referenced test for grades 4, 6 and 8 and a criterion-referenced instrument for 17-year-olds. The 17-year-old instrument contains selected NAEP items from citizenship, reading, social studies, science, math, writing, and career and occupational development, and is geared toward the basic skills. Following are assessments that used NAEP materials: 1975-76 -- reading, writing, math, citizenship/-social studies, science, career and occupational development; and 1977-78 -- reading, writing, math, citizenship/social studies, science, career and occupational development.

Texas

Texas replicated NAEP surveys in writing and citizenship for 9-, 13- and 17-year olds statewide in 1977-78. The intent was to follow NAEP's procedures as carefully as possible to permit comparisons between NAEP findings and performance levels within each of 20 education service center regions.

Washington

Washington uses NAEP items in math, reading and writing at grades 8 and 11. A norm-referenced test (CTBS) is administered at grade 4. All students are tested at the 4th grade level, but a state sample is taken at the 8th and 11th grades. Responses to the items are scored according to categories developed by NAEP, and results are given in comparison to national and western region scores. The State Council of Mathematics Teachers plans to examine the math findings on items common to the 8th and 11th grade assessments. Following are assessments that used NAEP materials: 1976-77 -- reading, math, writing; and 1978-79 -- reading, math, writing.

Replications by Other Countries

1976	Australia. Conducted a nationwide survey of reading, writing and arithmetic of 10- and 14-year-olds in more than 600 schools. Australia acknowledged considerable debt to National Assessment both in developing program objectives and in the use of several NAEP test items.

1976 — British Columbia. Conducted a pilot assessment of language arts for grades 4, 8 and 12. Some NAEP items were adapted for the survey, particularly at the fourth grade level.

Uses Other Than Replication

1979-80 — Massachusetts Speaking and Listening Project. The Massachusetts Department of Education is funding a one-year plan (1979-80) to develop instruments to assess basic skills in listening and speaking for secondary-level students. The effort is an extension of NAEP's recent feasibility study in speaking and listening and certain NAEP procedures were followed.

1979 — New York. The Research and Development Center for Occupational Education at City University of New York was awarded a grant from NIE to launch the first phase of a project that would result in a Spanish edition of selected items from National Assessment's 1973-74 career development survey. The purpose of the translation was to partial out the linguistic obstacles that might account for difficulties Hispanos encounter in answering NAEP exercises.

Local

1977 — Roanoke Rapids, North Carolina. Developed a mathematics assessment that included basic mathematical abilities it was felt that high school students needed to function as competent adults. Approximately 30% of the test items were taken from the NAEP mathematics assessment. Teachers, parents, mathematics educators and prospective employers were

	involved in making up the assessment.
1977	Brookline and Fitchburg, Massachusetts. U - NAEP writing items and scoring methodology in 1972-75 and 1977 to make comparisons to the nation and to identify curricular areas needing added emphasis.
1976	Cheyenne Mountain School District and Air Academy School District, Colorado Springs, Colorado. Both school systems conducted a replication of NAEP's writing assessment, scoring for mechanics and primary traits. Results were compared to National Assessment's findings.
1976	Lincoln, Nebraska. Conducted systemwide writing assessment of 4th, 7th and 10th graders using NAEP procedures and some NAEP items, modified to meet local needs. Results were used to revise curricula and develop training techniques for teachers in teaching writing.
1976	Bloomington, Minnesota. More than 4,900 students were involved in an assessment of writing that drew heavily upon methods and questions used in NAEP surveys. Fourth, 8th and 11th graders responded. The district wished to compare their results to NAEP's so a replication of much of the NAEP writing assessment was completed.
1976	Sioux Falls Independent School District No. 49-5, Sioux Falls, South Dakota. Received materials in reading, writing and

	mathematics to use in a study to determine whether or not to include some performance-based requirements with present graduation requirements.
1976	Edmunds School District, Washington. Utilized NAEP released exercises for assessment of their programs and comparison of their data to National data.

APPENDIX B

PEOPLE AND/OR ORGANIZATIONS CONTACTED BY ABT ASSOCIATES INC. AND BY THE EDUCATION COMMISSION OF THE STATES

People Interviewed by Abt Associates Inc.

Federal Legislative Branch

Allen P. Cissel, associate staff director, Subcommittee on Select Education, Committee on Education and Labor, U.S. House of Representatives.

William C. Clohan, Minority Staff, Committee on Education and Labor, U.S. House of Representatives.

Richard D. DiEugenio, Minority Staff, Committee on Education and Labor, U.S. House of Representatives.

Jean Frohlicker, counsel, Subcommittee on Education, Arts and the Humanities, Committee on Labor and Human Resources, U.S. Senate.

Polly Gault, Minority Staff, Committee on Labor and Human Resources, U.S. Senate.

Samuel W. Hunt, professional staff member, Committee on Appropriations, U.S. Senate.

John F. Jennings, counsel, Subcommittee on Elementary, Secondary and Vocational Education, Committee on Education and Labor, U.S. House of Representatives.

K. Forbis Jordan, Congressional Research Service.

David Kehl, professional staff member (minority), Committee on Appropriations, U.S. House of Representatives.

Robert F. Lyke, Congressional Research Service.

David J. Morse, Minority Staff, Committee on Labor and Human Resources, U.S. Senate.

David Mundel, assistant director, Congressional Budget Office.

Frederick Pfluger, staff assistant, Subcommittee on Labor, Health, Education and Welfare, Committee on Appropriations, U.S. House of Representatives.

Jennifer W. Vance, Minority Staff, Committee on Education and Labor, U.S. House of Representatives.

Thomas R. Wolanin, staff director, Subcommittee on Postsecondary Education, Committee on Education and Labor, U.S. House of Representatives.

Federal Agencies

Beatrice Berman, Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Education.

Marge Chandler, acting director of Post-Secondary Education and Vocational Education, National Center for Education Statistics, U.S. Department of Education.

John Evans, Office of Evaluation, U.S. Department of Education.

Iris Garfield, National Center for Education Statistics, U.S. Department of Education.

Alan Ginsberg, director of Elementary and Secondary Education, Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Education.

Mary Golladay, chief of Indicators Branch, National Center for Education Statistics, U.S. Department of Education.

Kenneth Hoyt, director of the Office of Career Education, U.S. Department of Education.

Shirley Jackson, director of the Basic Skills Program, U.S. Department of Education.

Marshall Smith, executive assistant to the secretary, U.S. Department of Education.

Michael Timpane, director, National Institute of Education, U.S. Department of Education.

Dustin Wilson, Office of Consumer Education, Bureau of School Improvement, U.S. Department of Education.

Carl Wisler, Office of Evaluation, U.S. Department of Education.

National Education Organizations

Wesley Apker, executive director, National Association of State Boards of Education.

Richard Duffy, U.S. Catholic Conference.

John Esty, president, National Association of Independent Schools.

David Florio, legislative liaison, American Education Research Association.

Jewell Gould, American Federation of Teachers.

Walter Haney, National Consortium on Testing (and the Huron Institute).

James Keefe, director of research, National Association of Secondary School Principals.

Edward Keller, National Association of Elementary School Principals.

Eugenia Kemble, special assistant to the president, American Federation of Teachers.

James Kirkpatrick, American Association of School Administrators.

William Pierce, Council of Chief State School Officers.

Sister Carleen Reck, executive director of Elementary Education Department, National Catholic Education Association.

Thomas Shannon, executive director, National School Boards Association.

Robert Smith, Council on American Private Education.

People Interviewed by the Education Commission of the States

Dr. Ray Billingsley, director, High School Improvement Project, Newton County Comprehensive High School, Covington, Ga.

Dr. Gerald Bracey, director, Research, Education and Testing, Virginia State Department of Education.

Dr. Nancy Burton, post-doctoral fellow, University of North Carolina.

Mr. Allan Duffer, regional supervisor, Research Triangle Institute, Raleigh, N.C.

Dr. Gordon Ensign, supervisor, Testing and Evaluation, Washington Department of Public Instruction.

Dr. Norris Harms, director, Project Synthesis, and evaluation specialist, School of Education, University of Colorado.

Dr. Allan Hartman, director, Research and Assessment, Massachusetts State Department of Education.

Dr. Thomas Innes, director of Educational Assessment, University of Tennessee, Knoxville, Tenn.

Dr. Ruth Jones, associate professor of Political Science, University of Missouri-St. Louis, St. Louis, Mo.

Dr. Stephen Jurs, professor and chairman of Department of Educational Research and Measurement, University of Toledo, Toledo, Ohio.

Mr. Karl Kurtz, National Conference of State Legislatures, Denver, Colo.

Dr. Meredith Larson, senior policy analyst, Stanford Research Institute International.

Dr. William McMillan, director, Minnesota Statewide Assessment, Minnesota Department of Education.

Dr. James McPartland, co-director, Center for Social Organization of Schools, John Hopkins University, Baltimore, Md.

Dr. Donald Miller, group director, School Training and Policies Group, CEMREL, St. Louis, Mo.

Dr. David Nelson, coordinator, Planning and Evaluation, Utah State Department of Education.

Mr. Glen Newkirk, National Conference of State Legislatures, Denver, Colo.

Dr. Edward Reidy, director of Project Competency, Fitchburg Public Schools, Fitchburg, Mass.

Dr. Robert Reys, professor of Mathematics Education, University of Missouri.

Dr. Edward Roeber, supervisor, Michigan Educational Assessment Project, Michigan State Department of Education.

Dr. Bruce Tone, editorial assistant to the President, International Reading Association.

Dr. Herbert Walberg, research professor of Urban Education, University of Illinois at Chicago Circle.

Dr. Celeste Woodley, curriculum specialist, Boulder Valley Schools, Boulder, Colorado; member, Board of Directors, National Council for the Social Studies.

BIBLIOGRAPHY

Anderson R. et al. "Toward a Methodology for Secondary Analysis of Data From the National Assessment of Educational Progress" (unpublished draft). Minneapolis, Minn.: University of Minnesota, April 1980.

Armor, D.J. "School and Family Effects on Black and White Achievement: A Reexamination of the USOE Data," in On Equality of Equal Opportunity. New York: Random House, 1972.

Burton, N. "Assessment as Exploratory Research: A Theoretical Overview." Paper presented at the annual meeting of the American Educational Research Association, 1978. ERIC no. ED 160 646.

Cohen, David K. "Politics and Research," Review of Educational Research, vol. 40, no. 2, 1970.

Cohen, David K. and Michael S. Garet. "Reforming Educational Policy With Applied Social Research," Harvard Educational Review, vol. 45, no. 1, February 1975.

Dyer, H. "Some Thoughts About Future Studies," in On Equality of Educational Opportunity. New York: Random House, 1972.

General Accounting Office, The National Assessment of Educational Progress: Its Results Need To Be Made More Useful. The Comptroller General of the United States, 1976.

Mosteller, F. and D. Moynihan. On Equality of Educational Opportunity. New York: Random House, 1972.

Patton, Michael Q. Utilization-Focused Evaluation. Beverly Hills, Calif.: Sage Publications, 1978.

Weiss, Carol H. "Introduction," in Using Social Research in Public Policy Making, ed. Carol H. Weiss. Lexington, Mass.: Lexington Books/D.C. Heath, 1977.

Welch, W.W. et al. "The Effects of Schooling on Math Achievement" (mimeographed copy). Minneapolis, Minn., 1980.